

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 56.4 Seconds

(without alignments)

842.900 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTSFSPFLFQEDKR.....LSHVQDCSTPGMNSPSPKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

- 1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2348	100.0	428	20 AAW73963	Bos taurus sialida
2	1916.5	81.6	428	20 AAW73964	Human sialidase pr
3	1916.5	81.6	461	22 AAB93971	Human protein sequ
4	692	29.5	16	AARG2547	CHO cell sialidase
5	295	12.6	245	22 AAU22595	Novel human colon
6	295	12.6	245	22 AAU22595	Human digestive sy
7	255.5	10.9	415	19 AAW92228	Human lysosomal si
8	213	9.1	151	22 AAW90545	Human immune/haema
9	192.5	8.2	492	20 AAY34503	Porphorymonas ging
10	192.5	8.2	540	20 AAY34376	Porphorymonas ging
11	176.5	7.5	516	22 AAU55185	Propionibacterium

12	154.5	6.6	544	17	AAW02207	Bacteroides fragil
13	146	6.2	396	15	AAAR47061	Bacteroides fragil
14	140.5	6.0	683	22	AAU03632	Group B Streptococ
15	138	5.9	248	21	AAB53418	Human colon cancer
16	137.5	5.9	394	22	AAG31465	C glutamicum prote
17	136	5.8	795	22	AAU45515	Propionibacterium
18	110	4.7	51	16	AARG2556	Tryptic peptide no
19	108.5	4.6	642	20	AAU01541	Alpha(2-3) trans-s
20	106.5	4.5	1060	20	AAU01540	Trypanosoma cruzi
21	104.5	4.5	763	22	ABG00868	Novel human diagn
22	102	4.3	680	20	AAU24320	Mouse dephosphoryl
23	99	4.2	567	22	AAAB8658	Human polypeptide
24	98	4.2	931	20	AAW96253	Mouse semaphorin r
25	97	4.1	925	20	AAW96308	Neuropilin-2. Rat
26	94.5	4.0	596	22	AAW40444	Human polypeptide
27	94.5	4.0	2820	22	ABG63296	Drosophila melanog
28	94	4.0	209	22	AAW66017	Propionibacterium
29	93	4.0	244	22	AAB79141	Corynebacterium gl
30	93	4.0	244	22	AAB79142	Corynebacterium gl
31	92	3.9	479	22	AAG75222	Amino acid sequenc
32	92	3.9	567	21	AAU93988	Amino acid sequenc
33	92	3.9	628	22	AAG65805	Human leucine-rich
34	92	3.9	628	22	AAB84469	Amino acid sequenc
35	92	3.9	628	22	AAG67523	Amino acid sequenc
36	92	3.9	636	22	AAU32870	Novel human secret
37	92	3.9	909	20	AAW96249	Rat semaphorin rec
38	91.5	3.9	782	18	AAW19764	Her2-GM-CSF immuno
39	91.5	3.9	926	20	AAW96252	Mouse semaphorin r
40	90.5	3.9	529	22	AAU65308	Propionibacterium
41	90	3.8	66	22	ABBI7728	Human nervous syst
42	90	3.8	380	22	ABG07024	Novel human diagn
43	89.5	3.8	1246	20	AAU06296	Human transcriptio
44	89	3.8	493	22	ABG63746	Drosophila melanog
45	89	3.8	1572	18	AAW27160	Mouse receptor ME2

#### ALIGNMENTS

RESULT 1

AAW73963

ID AAW73963 standard; Protein: 428 AA.

XX AC AAW73963;

XX AC AAW73963;

DT 29-APR-1999 (first entry)

XX Bos taurus sialidase protein sequence.

DE Membrane-associated ganglioside sialidase; sialic acid residue removal;

KW Carbohydrate structure determination; cancer; gene therapy;

XX brain disease; cow.

OS Bos taurus.

XX WO9853052-A1.

XX PD 26-NOV-1998.

XX PF 11-MAY-1998; 98WO-JP02072.

XX PR 22-MAY-1997; 97JP-0132174.

XX PA (MIYA-) MIYAGI-KEN.

PI Miyagi T, Wada T, Yoshikawa Y;

XX WPI: 1999-095226/08.

XX DR N-PSDB; AAX01559.

XX PT Membrane-associated ganglioside sialidase - useful in the

PT investigation of carbohydrate structures and for gene therapy of

PT brain disease



Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
Homo sapiens.  
EP1074617-A2.  
07-FEB-2001.  
28-JUL-2000; 2000EP-0116126.  
29-JUL-1999; 99JP-0248036.  
27-AUG-1999; 99JP-0300253.  
11-JAN-2000; 2000JP-018776.  
02-MAY-2000; 2000JP-0183767.  
09-JUN-2000; 2000JP-0241899.  
(HELI-) HELIX RES INST.  
Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
WPI; 2001-318749/34.  
Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -  
Claim 8; SEQ ID 14038; 2537pp + CD ROM; English.  
The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.  
Sequence 461 AA;  
Query Match 81.6%; Score 1916.5; DB 22; Length 461;  
Best Local Similarity 82.4%; Pred. No. 2.8e-168;  
Matches 352; Conservative 32; Mismatches 40; Indels 3; Gaps 2;  
QY 1 MEEVTSFSFSLFQBEDKRGVYTRIPALYVPPAHTFLFAEKRSSKDEDAHLHLVLR 60  
DB 34 meevttesfnspfrqgdggicyrpallyipthtflfaekstrrddalhlvlr 93  
QY 61 GLRTGQSVQHEPLKSLMKATLPGRHTMPCPVWERSGYVLFICVQGHVTERQIMSG 120  
DB 94 glrlgqlvqgplkpmeatlpghrtmnpvcvweqsgcvflfcvgrhvterqivsg 153  
QY 121 RNPALRCFCSQDAGYSWSVDRLTEEVIGVEVTHWATEAVGPGCHGJQLQSGRLITPAYA 180  
DB 154 rnaaricflqsdagcswsevdteevigseikhwatfavgpghqlqsggrlvipayt 213  
QY 181 YYIPFWFCFRLPYRAPHSLMIYSDDLGTATWHHRLIKPMVTVECEVAEVIKAGHPVL 240

DB 214 yyipswffcfqlpcktrphslmiysddlgvtwhhgrlirpmvtvecevaeavtgraghpvl 273  
QY 241 YCSARTPNRHRAEALSIDHGCECFQKPVLSHQLCEPHGCGSVVSECPLTIPGCGQDLAG 300  
DB 274 ycsartpnrcraealsthdgegfrlalsrqlcepphgcgsvvsfrplphrcqdssg 333  
QY 301 EDAPAIQOSPLLCSSVRPEAGTLESWLLYSHPNKKRRVDLGIYLNQSPLEAACWSR 360  
DB 334 kdaptlqgss-pgsslrleeaagtpsewllshphtarkrvdglylnatpleaacwv 392  
QY 361 PWILHCGPCGYSDLAALALENEGLFCCLFECCTKQECBOIAFRFLTDRILSHVQDCSTPG 420  
DB 393 pwilhcgpcgysdlaaleeeglfccfctkqecqiafrlthreilshlgdctspg 452  
QY 421 MNSEPSK 427  
DB 453 rn--psg 457  
RESULT 4  
ID AAR62547 standard; Protein; 379 AA.  
XX AAR62547;  
AC AAR62547;  
XX 09-AUG-1995 (first entry)  
DT 09-AUG-1995 (first entry)  
XX CHO cell stialdase.  
DE CHO cell stialdase.  
XX Sialidase; sialic acid; Chinese Hamster ovary.  
XX Crictetus crictetus.  
OS Crictetus crictetus.  
FH Key Location/Qualifiers  
FT Peptide 11..29  
FT Peptide 112..136  
FT Peptide 45..75  
FT Peptide /label= PCR 14/17 probe  
FT Peptide 139..169  
FT Peptide 172..201  
FT Peptide 208..218  
FT Peptide 244..263  
FT Peptide 284..335  
FT Peptide 231..237  
PN WO9426908-A.  
XX 24-NOV-1994.  
PD 24-NOV-1994.  
XX 17-MAY-1994; 94MO-US05471.  
PF 17-MAY-1994; 94MO-US05471.  
XX 17-MAY-1993; 93US-0062586.  
PR 25-JAN-1994; 94US-0187327.  
XX (GETH ) GENENTECH INC.  
PA Sliwkowski MB, Warner TG;  
PI Sliwkowski MB, Warner TG;  
XX WPI; 1995-006802/01.  
DR N-PSDB; AAQ75442.  
XX Recombinant cell line defective for sialidase expression - useful for prodn. of glyco:protein(s), also new CHO cell sialidase and related DNA etc., for treating inflammation, lung diseases, etc.  
PT Disclosure; Figure 10; 64pp; English.  
PS Purified sialidase from CHO cells was treated with various proteases and the resulting peptides sequenced (see, for example, AAR62565). The lysine C derived peptide LC18 (see AAR62565) is a composite of tryptic peptides TP14 and TP17 and is referred to as TP14/17. A PCR 14/17 probe was prepd. and used





PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0251989.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Harash SC, Ruben SM;  
XX  
XX WPI: 2001-502630/55.  
XX N-PSDB; AAK88342.  
XX  
XX Polynucleotides encoding digestive system antigens, useful for  
XX diagnosing, treating, preventing and/or prognosing disorders of the  
XX digestive system, particularly cancer and cancer metastases -  
PS Claim 11; SEQ ID NO 1918; 986pp; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
XX number of human digestive system antigens. These can be used in the  
XX diagnosis, treatment and prevention of digestive system disorders,  
XX including cancer, Meckel's diverticulum, bacterial or parasitic  
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or  
XX ulcerative colitis. The present sequence is a digestive system antigen of  
XX the invention.  
XX  
XX Sequence 245 AA;

Query Match 12.6%; Score 295; DB 22; Length 245;  
Best Local Similarity 62.1%; Pred. No. 5.4e-19;  
Matches 54; Conservative 9; Mismatches 24; Indels 0; Gaps 0;



PR	22-AUG-2000;	2000US-02286860;
PR	22-AUG-2000;	2000US-02271882;
PR	23-AUG-2000;	2000US-02271702;
PR	30-AUG-2000;	2000US-02289294;
PR	01-SEP-2000;	2000US-02292887;
PR	01-SEP-2000;	2000US-02293143;
PR	01-SEP-2000;	2000US-02293344;
PR	01-SEP-2000;	2000US-02293345;
PR	05-SEP-2000;	2000US-02295019;
PR	05-SEP-2000;	2000US-02295193;
PR	06-SEP-2000;	2000US-02304337;
PR	06-SEP-2000;	2000US-02304338;
PR	08-SEP-2000;	2000US-02312423;
PR	08-SEP-2000;	2000US-02312433;
PR	08-SEP-2000;	2000US-02312434;
PR	08-SEP-2000;	2000US-02312444;
PR	08-SEP-2000;	2000US-02314113;
PR	08-SEP-2000;	2000US-02314114;
PR	08-SEP-2000;	2000US-02320800;
PR	08-SEP-2000;	2000US-02320808;
PR	12-SEP-2000;	2000US-02319667;
PR	14-SEP-2000;	2000US-02333997;
PR	14-SEP-2000;	2000US-02333998;
PR	14-SEP-2000;	2000US-02333999;
PR	14-SEP-2000;	2000US-02334200;
PR	14-SEP-2000;	2000US-02334201;
PR	14-SEP-2000;	2000US-02330633;
PR	14-SEP-2000;	2000US-02330634;
PR	14-SEP-2000;	2000US-02330635;
PR	21-SEP-2000;	2000US-02342223;
PR	21-SEP-2000;	2000US-02342274;
PR	25-SEP-2000;	2000US-02349977;
PR	25-SEP-2000;	2000US-02349998;
PR	26-SEP-2000;	2000US-02354884;
PR	27-SEP-2000;	2000US-02358834;
PR	27-SEP-2000;	2000US-02358836;
PR	29-SEP-2000;	2000US-02363627;
PR	29-SEP-2000;	2000US-02363637;
PR	29-SEP-2000;	2000US-02363688;
PR	29-SEP-2000;	2000US-02363670;
PR	29-SEP-2000;	2000US-02363700;
PR	02-OCT-2000;	2000US-02368023;
PR	02-OCT-2000;	2000US-02370377;
PR	02-OCT-2000;	2000US-02370388;
PR	02-OCT-2000;	2000US-02370399;
PR	02-OCT-2000;	2000US-02370400;
PR	13-OCT-2000;	2000US-02399335;
PR	13-OCT-2000;	2000US-02399337;
PR	20-OCT-2000;	2000US-02409660;
PR	20-OCT-2000;	2000US-02412221;
PR	20-OCT-2000;	2000US-02417885;
PR	20-OCT-2000;	2000US-02417886;
PR	20-OCT-2000;	2000US-02417887;
PR	20-OCT-2000;	2000US-02418008;
PR	20-OCT-2000;	2000US-02418009;
PR	20-OCT-2000;	2000US-02418266;
PR	01-NOV-2000;	2000US-02446177;
PR	08-NOV-2000;	2000US-02464774;
PR	08-NOV-2000;	2000US-02464775;
PR	08-NOV-2000;	2000US-02464776;
PR	08-NOV-2000;	2000US-02464777;
PR	08-NOV-2000;	2000US-02464778;
PR	08-NOV-2000;	2000US-02465323;
PR	08-NOV-2000;	2000US-02465245;
PR	08-NOV-2000;	2000US-02465246;
PR	08-NOV-2000;	2000US-02465247;
PR	08-NOV-2000;	2000US-02465328;
PR	08-NOV-2000;	2000US-02465329;
PR	08-NOV-2000;	2000US-02465332;
PR	08-NOV-2000;	2000US-02465609;
PR	08-NOV-2000;	2000US-02466100;
PR	08-NOV-2000;	2000US-02466111;
PR	08-NOV-2000;	2000US-02466133;
PR	17-NOV-2000;	2000US-02492077;





```
CC especially gingivitis.
XX
SQ Sequence 540 AA;

Query Match      8.2%; Score 192.5; DB 20; Length 540;
Best Local Similarity 22.7%; Pred. No. 4.6e-09;
Matches 93; Conservative 60; Mismatches 164; Indels 93; Gaps 16;

QY 13 LFOQEDKGVNRYRIPALIVYPPAHTFLFAFAEKRSSKDE--DALHLVLRGLRTGQSQVOW 70
Db 196 lfvpdgggsrnyripailktang-tliamadrkynqtdlpedidivmrtdgkgs--w 252
QY 71 EPLSLMKATLPGHRTMPCPVWERKSGYVLFYFCVGOGHTERQQIMSGRNPALRCFIC 130
Db 253 sdprliivgegrnh-gfgdvalvtqagkllmifv---ggvlgwqs--tpdrpqrtyise 306
QY 131 SODAGYSWSDVRDLTEEVIG-----PEVTHWATFVAGFGHGIQOSGRLLIIPA-----Y 179
Db 307 srdegltwsprrdithfigkdcadpgrsrwlasfcasggglvlpsggrttfvaalresgq 366
QY 180 AYYIPFWFCFRLPYRAPHSLMIYSDDLGATW-----HHGRLIKPMVTVECEVAEVI 232
Db 367 eyvl-----nnyvlvsddgdgtwqlsdcayrgrdeaklsimpdgrvlmsi 411
QY 233 GKAGHPVLVCSARTPNRHRABALSIDHGECFKPKVLSHQLCEPPHGCQGSVVSFCPLRIP 292
Db 412 rnqg-----rqesqrffalsddglwtweraqkfegihdp--gcngam-----lqvk 456
QY 293 GGCQDLAGEDAPATQOSPLLCSSLVRPEPEAGTLESWLLYSHPTNKKRRVDGIYLNOSP 352
Db 457 rngrd-----qvlhslpdpdrrdgavylfdh- 484
QY 353 LEACWSRPWILHCGPCGYSDLAALENGLFCGLFCEGCKQECQIAFRL 402
Db 485 -vsgrwsapvvvngssaysdmtila-dgtigyfveesdeislvfirfvl 532

RESULT 11
AAU55185
ID AAU55185 standard; Protein; 516 AA.
XX
AC AAU55185;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #16081.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
XX
XX 02-JUN-2000; 2000US-208841P.
XX
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX
XX N-PSDB; AAS959568.
```

```
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Claim 3; SEQ ID No 16380; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 516 AA;

Query Match      7.5%; Score 176.5; DB 22; Length 516;
Best Local Similarity 24.1%; Pred. No. 1.3e-07;
Matches 95; Conservative 52; Mismatches 158; Indels 89; Gaps 19;

QY 17 EDKRGVYRIPALIVYPPAHTFLFAFAEKRSSKD-EDALHLVLRGLRTGQSQVQWEPLKS 75
Db 127 ehpegvfyrpalatasngdllasydlrpsagdapnpsivqrrsrdngrt--wgp-qt 183
QY 76 LMKATLPGHRTM---NPCVWERKSGYVYLFPI--CVOGHVTERRQIMSGRNPALRCFI- 129
Db 184 vihagtprrkvgysdpsylvdpatgrilnfhvksydrgfatse-----vgtdpdrhrvlh 239
QY 130 ---CSQDAGYSWSDVRDLTEEVIGPEVTHWATFVAGFGHGIQL---OSGRLLIIPAYAY 181
Db 240 aevststdnghgtwh-rditreitsdptr--trfvasggqiallhghagrliaq---- 292
QY 182 YIPFWFCFRLPYRAPHSLMIYSDDLGATWHHGLIKPMVTVECEVAEVIKAGHPVLY 241
Db 293 -----mtvrnsvggqagslysdhgitwhagnpvgmmnd-enkvvel-----sdgtlm 339
QY 242 CSARTPNR--HRABALSIDHGECFKPKVLSHQLCEPPHGCQGSVVSFCPLRIPGQCQDLA 299
Db 340 lnsrdaarsgrkvaysqdggltwgpkvlvddldtpnnaq----- 380
QY 300 GEDAPATQOSPLLCSSLVRPEPEAGTLESWLLYSHPTNKKRRVDGIYLNOSPLEAAC-W 358
Db 381 -----firaypnaragsakarillftnarnaternv-----gtlsvscdd 420
QY 359 SRPWILH---CGPCGYSDLAALENGLFCGLFE 388
Db 421 grtwshqtympegvgt-taavsdgdlgvlwe 453

RESULT 12
AAU02207
ID AAW02207 standard; Protein; 544 AA.
XX
AC AAW02207;
XX
XX 11-OCT-1996 (first entry)
XX
XX Bacteroides fragilis neuraminidase.
XX
```

KW Neureminidase; identification; diagnosis; pathogenic factor;  
 KW stabilisation; sialic acid containing substance.

OS Bacteroides fragilis.

PN JP08131174-A.

XX 28-MAY-1996.

XX 15-NOV-1994; 94JP-0280237.

XX 15-NOV-1994; 94JP-0280237.

XX (SAKA ) OTSUKA PHARM CO LTD.

XX WPI; 1996-303851/31.

XX N-PSDB; AAT36720.

XX Bacteroides fragilis neureminidase gene - useful for identification  
 PT and diagnosis of B. fragilis

XX Claim 2; Pages 9-11; 12pp; Japanese.

CC The present sequence is the B. fragilis neureminidase (NA),  
 CC which was extracted from B. fragilis using, e.g. the method of  
 CC Thomas (Berns, L.I. and Thomas, Jr., C.A. J. Mol. Biol., 11,  
 CC 476-490 (1965)). The whole or a part (pref. an EcoRI-Cla fragment)  
 CC of the NA gene can be used for the identification and diagnosis of  
 CC B. fragilis, a microorganism believed to have a NA a pathogenic  
 CC factor not considered to be present in other bacterial spp.. The  
 CC NA may be used to stabilise a sialic acid-contg. substance.

XX Sequence 544 AA;

Query Match 6.6%; Score 154.5; DB 17; Length 544;  
 Best Local Similarity 21.6%; Pred. No. 1.5e-05;  
 Matches 84; Conservative 59; Mismatches 157; Indels 89; Gaps 15;

QY 24 YRIPALIVPPAHTFLAFAEKRSKSD-EDALHLVLRGLRTGQSVQWE-----PLKSLMK 78

DB 203 frlpplvttnkgtllgvdyvrynsdvldqehvdvlglsrstdggt--wekmrlplafgef 260

QY 79 ATLPG--HRTMNPVNERKSGYVLFICVQGHVTER-----QQIMSGRNPRLCFICS 131

DB 261 gglpagngvgdpsilvdtktnvwwvaawthgmgngqrawwshpmdmhtaglvlaaks 320

QY 132 ODAGYSNDVRDLTEEVIGPEVTHMATFVAVGPHGHIQLOSGRLIIPAYAYIPIFWFFCFR 191

DB 321 tddgktwsapiniteqvkdp---wyflggpggrgltmsdgtlvfptqf----- 366

QY 192 LPYRAPHSLMIYSDDLGATWHGRLIKPMVTVECEVAEVIKAGHPVLYCSARTPNR-- 249

DB 367 idstrvnpagimyskdggknwkmhnyartn--tteaqvaev-----epgvlmnmrdnrg 420

QY 250 HRBAALSIDHGECFOKPVLSHOLCEPPHGCQSVVSPCLPIPGCCQDLAGEDAPAIQOS 309

DB 421 sravaitkdkgktwtehesark-----alipes 447

QY 310 PLLCSSLVRPEPEAGTLESWLLYSHTNKKRRVDLGYLNQSPLEAACWS----- 359

DB 448 vcmassliisvakdnvlgkdlifsnpttkgrynttikisid--ggvtspehqlildeg 505

QY 360 RPWILHCGPCGYSDLAALENEGLFGCLFE 388

DB 506 nnw-----gysclsmidketi-gliye 526

RESULT 13

AAR47061

ID AAR47061 standard; Protein: 396 AA.

XX AAR47061;

XX 03-OCT-1994 (first entry)

XX Bacteroides fragilis neureminidase.

XX Neureminidase; B. fragilis; detection.

XX Bacteroides fragilis.

XX JP06046865-A.

XX 22-FEB-1994.

XX 17-JUL-1991; 91JP-0269931.

XX 17-JUL-1991; 91JP-0269931.

XX (SAKA ) OTSUKA PHARM CO LTD.

XX WPI; 1994-097024/12.

XX N-PSDB; AAQ58660.

XX Neureminidase gene from Bacteroides fragilis - useful for prodn.  
 PT of neureminidase and detection of B. fragilis

XX Claim 2; Page 10-11; 14pp; Japanese.

CC The gene from B. fragilis YCH46 coding for neureminidase AAR47061 is  
 CC useful for recombinant production of the enzyme and for the  
 CC identification and diagnosis of B. fragilis infection.

XX Sequence 396 AA;

Query Match 6.2%; Score 146; DB 15; Length 396;  
 Best Local Similarity 21.5%; Pred. No. 5.9e-05;  
 Matches 84; Conservative 59; Mismatches 157; Indels 90; Gaps 16;

QY 24 YRIPALIVPPAHTFLAFAEKRSKSD-EDALHLVLRGLRTGQSVQWE-----PLKSLMK 78

DB 54 frlpplvttnkgtllgvdyvrynsdvldqehvdvlglsrstdggt--wekmrlplafgef 111

QY 79 ATLPG--HRTMNPVNERKSGYVLFICVQGHVTER-----QQIMSGRNPRLCFICS 131

DB 112 gglpagngvgdpsilvdtktnvwwvaawthgmgngqrawwshpmdmhtaglvlaaks 171

QY 132 ODAGYSNDVRDLTEEVIGPEVTHMATFVAVGPHGHIQLOSGRLIIPAYAYIPIFWFFCFR 190

DB 172 tddgktwsapiniteqvkdp---wyflggpggrgltmsdgtlvfptqf----- 218

QY 191 LPYRAPHSLMIYSDDLGATWHGRLIKPMVTVECEVAEVIKAGHPVLYCSARTPNR- 249

DB 219 -idstrvnpagimyskdggknwkmhnyartn--tteaqvaev-----epgvlmnmrdnrg 271

QY 250 -HRAEALSIDHGECFOKPVLSHOLCEPPHGCQSVVSPCLPIPGCCQDLAGEDAPAIQO 308

DB 272 gsraaitkdkgktwtehesark-----alpe 298

QY 309 SPLCSSLVRPEPEAGTLESWLLYSHTNKKRRVDLGYLNQSPLEAACWS----- 359

DB 299 svcmassliisvakdnvlgkdlifsnpttkgrynttikisid--ggvtspehqlilde 356

QY 360 -RPWILHCGPCGYSDLAALENEGLFGCLFE 388

DB 357 gnnw-----gysclsmidketi-gliye 378

RESULT 14

AAU03632

ID AAU03632 standard; Protein: 683 AA.

XX AAU03632;

DT 12-SEP-2001 (first entry)  
XX Group B Streptococcus antigenic protein, ID-107.  
DE  
XX Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;  
KW meningitis; neonate; antigenic; vaccine; infection; genital tract;  
KW capsid polysaccharide vaccination.  
XX  
OS Streptococcus agalactiae.  
XX  
XX  
PN WO200132882-A2.  
XX  
PD 10-MAY-2001.  
XX  
XX 07-SEP-2000; 2000WO-GB03437.  
PF  
XX 07-SEP-1999; 99GB-0021125.  
PR  
XX (MICR-) MICROBIAL TECHNIQS LTD.  
PA  
XX Le Page RWF, Wells JW, Hanniffy SB;  
PI  
XX WPI; 2001-316444/33.  
DR N-PSDB; AAS07049.  
XX  
XX New polypeptides derived from Streptococcus agalactiae are useful to  
PT provide detection of, and vaccination against, Group B Streptococcus  
PT infections, particularly to prevent infection in neonates.  
XX  
XX Claim 1; Fig 1; 178pp; English.  
PS  
XX AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus  
CC agalactiae) amino acid sequences of the invention. S. agalactiae is an  
CC encapsulated bacterium which is a major pathogen of humans causing sepsis  
CC and meningitis in neonates as well as adults. The S. agalactiae antigenic  
CC polypeptides are used to vaccinate against Group B Streptococcus  
CC infections, particularly to prevent infection in new born children  
CC arising from the maternal genital tract. An immunogenic composition is  
CC useful in the preparation of a medicament for the treatment or  
CC prophylaxis of Group B Streptococcus infection. The invention does not  
CC have the disadvantages of varied response rate associated with prior art  
CC capsid polysaccharide vaccination against Group B Streptococcus.  
XX  
XX Sequence 683 AA;  
SQ  
Query Match 6.0%; Score 140.5; DB 22; Length 683;  
Best Local Similarity 22.1%; Pred. No. 0.0004;  
Matches 77; Conservative 45; Mismatches 109; Indels 117; Gaps 15;  
QY 84 HRTMPCPVWERKSGVYVLFICVQGHVTERQIMSGRNPALCFICSDAGYSWSDVRD 143  
Db 388 hkt---spflakssyywnsy-----sdddgtrwsprd 418  
QY 144 LTEEIGPEVTHWATFVPGHGHIQLQ----SGRLIIPAYAYIPFWFFCRLPYRAPH 199  
Db 419 itasl---rqgmkfkgipgkgivlkwpghagriiipaystn-----wkshlrgsq 468  
QY 200 SLMIYSDDLGATWHHGRLLKP-----WVTVE-----CEVAEIVGKAGHPVLVC 242  
Db 469 rliysddhgktwhtgkavndnrilnsgekihsaltmdnkkeqntesvpvqlkngdiklfm 528  
QY 243 SARTPNRRAEALSDHGECFQKPVLSHQLCEPPHCQSGSVVSFCPLIPGGCODLAGED 302  
Db 529 rnltnle---vatskggetqwnhvkryk-----555  
QY 303 APAIQSPILCSVRPEPEAGTILSESWLLSHYPTNKKRRVDLG-----IYLINQSPLEAACW 358  
Db 556 --ehdayqlsairfhd---kkeyillvnangpgkcdqgyarlaqvrn-----g 603  
QY 359 SRPWILH-----CGPCGYSDLALENGLCFLFCGCTKQCEQIAFRL 402  
Db 604 sfkwyhhhiqdsfaynsqvqlnnd-qfgviye---hrekhnstfl 647

RESULT 15  
AAB53418  
ID AAB53418 standard; Protein; 248 AA.  
XX  
AC AAB53418;  
XX  
DT 09-MAR-2001 (first entry)  
XX  
DE Human colon cancer antigen protein sequence SEQ ID NO:958.  
XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW identification; cytostatic; cardioactive; neuroprotective; vulnary;  
KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;  
KW neural disorder; immune system disorder; muscular disorder;  
KW reproductive disorder; gastrointestinal disorder; renal disorder;  
KW infectious disease; cardiovascular disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200055351-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 08-MAR-2000; 2000WO-US05883.  
XX  
PR 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX WPI; 2000-587534/55.  
DR N-PSDB; AAC98175.  
XX  
XX Colon cancer associated gene sequences, referred to as colon cancer  
PT antigens, useful for the treatment, prevention, and diagnosis of colon  
PT disorders such as colon cancer.  
XX  
XX Claim 1; Page 1523; 2104pp; English.  
XX  
XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The  
CC human colon cancer antigens can have cytostatic, cardioactive, muscular,  
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and  
CC vulnary, nephrotropic, antiinfective and antibacterial activities, and  
CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
CC proteins and antibodies to the proteins are useful for the prevention,  
CC treatment and diagnosis of colon disorders, such as colon cancer. The  
CC polynucleotides may be used in diagnostics and research, such as for  
CC chromosome identification, and as hybridisation probes. The proteins  
CC may also be used to prevent diseases such as neural disorders, immune  
CC system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, wounds, renal disorders, infectious  
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
CC AAB54007 represent sequences used in the exemplification of the present  
CC invention.  
XX  
XX Sequence 248 AA;  
SQ  
Query Match 5.9%; Score 138; DB 21; Length 248;  
Best Local Similarity 25.9%; Pred. No. 0.00017;  
Matches 74; Conservative 26; Mismatches 80; Indels 106; Gaps 15;  
QY 149 IGPEVTHWATFVPGHGHIQLQ----SGRLIIPAYAYIPFWFFCRLPYRAPHSLMIY 204  
Db 24 igtv-----fapgsqgikqreprkgrllvcgthgtdgvc-----ll 65  
QY 205 SDDLGTATWHHGRLLKPMTVTECEVAEIVGKAGHPVLVCYCSARTPNRRAEALSDHGECFQ 264  
||| |||:|:|  
|:|:|:|

Db 66 sddhgaswrygs-----gvsgip-----yggpkq 89  
Qy 265 KPVLSHQICEPPHGCQGSV-----SFCPLIPGCCQDLAGEDAPAIQQSPLLCSSV 316  
Db 90 endfnpdecqpyelpdgsvvlnarnqnnyhchriv-----lrsyda-----cdtl 135  
Qy 317 RP-----EPE-----AGTL--SESLLYSHTNKKRRVDLGIYLNOSPLEAACWSRPW 362  
Db 136 rprdvtfdpelvdpvvaagavvtssglvffsnpahpefrvnl--tlrwsfngtswrket 193  
Qy 363 I-LHCGPCGYSDLAALENEGLFG-----CLFECGTKQCEQIA 399  
Db 194 vqlwpgpsgysslatleg-smdgeeqapqlvlyvekgrnhytesis 238

Search completed: October 7, 2002, 15:00:52  
Job time: 134 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 32.29 Seconds  
(without alignments)  
1273.652 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTCSEFSSPLFQEDKR.....LSHVQGDCTPGMNSPPSKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1518	64.7	418	2 JC7588	exo-alpha-sialidas
2	708.5	30.2	385	2 JC7751	exo-alpha-sialidas
3	705	30.0	379	2 A49679	exo-alpha-sialidas
4	692	29.5	379	2 A54961	sialidase, cytosol
5	259	11.0	648	2 T34999	probable neuramida
6	201.5	8.6	647	2 A5244	exo-alpha-sialidas
7	197	8.4	748	2 T37097	probable secreted
8	186	7.9	1014	1 NMCLSS	exo-alpha-sialidas
9	172	7.3	694	2 I40866	exo-alpha-sialidas
10	169	7.2	901	2 A49227	sialidase - Actino
11	154	6.6	913	2 S20590	exo-alpha-sialidas
12	149	6.3	1035	2 T30287	exo-alpha-sialidas
13	149	6.3	1035	2 G38063	exo-alpha-sialidas
14	146	6.2	545	2 JC2500	exo-alpha-sialidas
15	145.5	6.2	404	2 A37234	exo-alpha-sialidas
16	128	5.5	382	2 S01339	exo-alpha-sialidas
17	125	5.3	740	2 G35153	neuraminidase, pro
18	120	5.1	382	2 S32148	exo-alpha-sialidas
19	117.5	5.0	773	2 JF0387	exo-alpha-sialidas
20	107	4.6	1548	2 S34533	serine proteinase
21	106.5	4.5	879	2 S23006	shed acute-phase a
22	105	4.5	781	2 A43866	neuraminidase - Vi
23	105	4.5	807	2 E82158	neuraminidase VC17
24	102	4.3	680	2 JC5895	killer cell inhibi
25	101	4.3	376	1 NNEBST	exo-alpha-sialidas
26	100	4.3	2767	1 UTHU	thyroglobulin prec
27	95.5	4.1	389	2 T08135	sedoheptulose-bisp
28	94.5	4.0	1216	2 A55620	apical endosomal p
29	91	3.9	697	2 E95196	neuraminidase B [1

RESULT 1

JC7588

exo-alpha-sialidase (EC 3.2.1.18) - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001

C:Accession: JC7588

R:Hasegawa, T.; Carnero, C.F.; Wada, T.; Itoyama, Y.; Miyagi, T.

Biochem. Biophys. Res. Commun. 280, 726-732, 2001

A:Title: Differential expression of three sialidase genes in rat development.

A:Reference number: JC7588; MUID:21092669; PMID:11162581

A:Contents: Brain

A:Accession: JC7588

A:Molecule type: mRNA

A:Residues: 1-418 <HAS>

A:Cross-references: DDBJ:AB026841

C:Comment: This enzyme is involved in metabolic regulation and cellular functions, in

C:Keywords: glycosidase; hydrolase

F:24-27/Region: consensus motif YRIP

F:131-138/Region: consensus Asp-box, three copies #status predicted

F:173-195/Domain: hydrophobic #status predicted <HYD>

F:203-210/Region: consensus Asp-box, three copies #status predicted

F:254-261/Region: consensus Asp-box, three copies #status predicted

Query Match 64.7%; Score 1518; DB 2; Length 418;  
Best Local Similarity 65.9%; Pred. No. 7e-123;  
Matches 278; Conservative 58; Mismatches 82; Indels 4; Gaps 2;

QY	1	MEEVTCSEFSSPLFQEDKRGVYTRIPALYYPPAHTFLAFAEKRSSKDEALHLVLR	60
DB	1	MEEVSSCLSRSLFQEEQNRTYTRIPALLYIPPTHTFLAFAEKRTSSHDEAVLVFRR	60
QY	61	GLRTQSQVQWEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFVFCVQGHVTERQIMSG	120
DB	61	GVWKCSVEMGQQQLMEATLPGHRTMPCPVWERNTGRVYLFVFCVQGHVSRWQLWG	120
QY	121	RNPALCTCSODAGYSNVDRLTEEVIGPEVTHAVFGPHGCIQLQSGRLIIPAYA	180
DB	121	RNARLCFLSYSDSCSGWGEVKDLTEEVVSGEMKHWATFVAVGPHGCIQLQSGRLIIPAYA	180
QY	181	YIYPMFCCFLPYRARPHSLMIYSDDLGATWHGRLTKPMYTVCEVAEVTGKAGHPVL	240
DB	181	YLISCFWLCF--PCSVKPHSLMFYSDDLGVTWHCKFKIPQVTGECQVAEVPKAGNLVL	238
QY	241	YCSARTPNRHRAEALSDHGECFQKPVLSHQLCPEPHGCGQSVSWFCEIPGCGQDLAG	300
DB	239	YCSARTPNKFAEAFSTDGDCFKQKTLNQQLCEPRGCGQSVSVTRPLKMPYTCODSSG	298
QY	301	EDAPAIQSPPLICSSVRPEPEAGTISESWLLYSHTPTNKKRRVDLGIYLNQSPLEACWSR	360
DB	299	KDVPSTQKPLMDRSLEVEEGAGPSTGLLYSHPTNKKRNMLGIYVYNNQNPLEVYNWSR	358

hypothetical prote  
surface antigen gp  
pyruvate flavodoxi  
conserved hypothet  
Photinus-luciferin  
Photinus-luciferin  
BFLF1 protein - hu  
phage-related prot  
killer cell inhibi  
exo-alpha-sialidas  
hypothetical prote  
hypothetical prote  
hypothetical prote  
carboxypeptidase-r  
C4BP alpha chain p

Qy	361	PWILHCPCGYSDLAALENEGLFCGLFCGCTKQECQIAFRFLTDRREILSHVQDGCSTG	420
		:             :                 :                 :	
Db	359	PWILNRPGSYSDLAVVGGGLFACLFECGERHEDEKIDFLCFSDQEVLS--CDDCTSPS	416
Qy	421	MN	422
Db	417	SN	418

```

RESULT      2
JC7751
exo-alpha-sialidase (EC 3.2.1.18) - mouse
N:Alternate names: thymic sialidase; neuraminidase; sialidase
C:Species: Mus musculus (house mouse)
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 14-Dec-2001
C:Accession: JC7751
R:Kotani, K.; Kuroiwa, A.; Saito, T.; Matsuda, Y.; Koda, T.; Kijimoto-Ochiai, S.
Biochem. Biophys. Res. Commun. 286, 250-258, 2001
A:Title: Cloning, chromosomal mapping, and characteristic 5'-UTR sequence of murine cyto
A:Reference number: JC7751; MUID:21391602; PMID:11500029
A:Accession: JC7751
A:Molecule type: mRNA
A:Residues: 1-385 <KOT>
A:CROSS-references: DDBJ:AB048604
C:Comment: This enzyme, a cytosolic type of sialidase, has high soluble sialidase activi
C:Genetics:
A:Gene: mts
A:Map position: 1d, distal part
C:Keywords: hydrolase; glycosidase

```

Query Match 30.2%; Score 708.5; DB 2; Length 385;  
Best Local Similarity 39.2%; Pred. No. 3.2e-53;  
Matches 166; Conservative 67; Mismatches 121; Indels 69; Gaps 15;

QY	58	LRLGL--RTGQS	VQWEPKLSMKATLPUGHRTWNPCDWPBERKSGYYVLFPPFCVQGHVTERQ	115
Dd	60	LRRGSYNEATNRVKWQP	EVEVTQAQLEGRHSNMPCLDYDKTKTLFTFIAPVRSEHH	119
QY	116	QIMSGRNPALRCFICS	SODAGYSWSVDRLTEEVIGPEVTHMATEFAVGPGHGLOLO--SGR	173
Dd	120	QLHTKNVTRLCCVS	TDHGRTWSPJQDLTETTGSTHOEWATFAVGPHSLQLRNPAQS	179
QY	174	LIIPAYAY-	-----IPFWFCFLPRPARRPSLMIYSDDLGATWHHGRLIKPMWTVE	225
Dd	180	LLVPAYAYRKLHPA	OAKPTPF-AFCF-----ISLDHGHTWKLGNFVAEN-SLE	224
QY	226	CEVAEVIKGAGHPVL	CSARTPNRHRAEAALSDHCFCOKPVLSHOLCEBPHGCOSVVVS	285
Dd	225	CQVAEY-GTGAQRMY	VLNARSELGARVQAQSPNDGLFDQDNRVVSKLVBPHPCHGSVVYA	283
QY	286	FCPLEIPGGQDLAGE	DAPAIQQSPLLCSSVRPEPAGTLESWLLYSHPTNKKRRVDLG	345
Dd	284	F-----	-----HNPI-----SKPHA---LDTWLTYHTDTSRNRNTNIG	313
QY	346	IYLNQSPLEAACWRSR	WIILHCGPCGYSDLAAL-----ENEGFLGCFECGTQKECQIAFR	401
Dd	314	VYLNQMLPDTAWSE	PTELLAMGICAYSIDLQNNGQGDPGSPQGCLEYSG--NYEEIIFL	370
QY	402	LFT	404	
Dd	371	IFT	373	

RESULT 3  
A49679  
exo-alpha-sialidase (EC 3.2.1.18), skeletal muscle, cytosolic - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-May-1995 #sequence\_revision 26-May-1995 #text\_change 05-

C:Accession: A49679; A57402  
R: Miyagi, T.; Konno, K.; Emori, Y.; Kawasaki, H.; Suzuki, K.; Yasui, A.; Tsuiki, S.  
J. Biol. Chem. 268, 26435-26440, 1993  
A: Title: Molecular cloning and expression of cDNA encoding rat skeletal muscle cytosolic  
A: Reference number: A49679; MUID: 94075331  
A: Accession: A49679  
A: Status: preliminary  
A: Molecule type: mRNA  
A: Residues: 1-379 <Miy>  
A: Cross-references: GB:D16300; NID: g434794; PIDN: BAA03805.1; PID: d1004321; PID: g43479  
R: Sato, K.; Miyagi, T.  
Glycobiology 5, 511-516, 1995  
A: Title: Genomic organization and the 5'-upstream sequence of the rat cytosolic sialyl  
A: Reference number: A57402; MUID: 96089543  
A: Accession: A57402  
A: Status: preliminary; not compared with conceptual translation  
A: Molecule type: DNA  
A: Residues: 1-67 <SAT>  
A: Cross-references: GB:D50606; NID: g1064833; PIDN: BAA09169.1; PID: d1009807; PID: g1311  
C: Keywords: cytosol; glycosidase; hydrolase; skeletal muscle

Query Match	30.0%	Score	705;	DB 2;	Length	379;			
Best Local Similarity	39.5%;	Pred. No.	6.3e-53;						
Matches	167;	Conservative	60;	Mismatches	126;	Indels	70;	Gaps	
Qy	12	PLFOQE--DKRGVYRIPALIVYPPAHTLAFAPKRSKDDDALHLVLRGLRTGQS- 67							
Db	5	PVLOKETLFTVEYAYRIPALLYKKQKTLTLLAFAPKRASTRDDEHAELIVLRGSGYNATN 64							
Qy	68	-VOMEPLKSLMKATLPGHRTMPCPVWERKSGVYVLFICVOGHVTEROOINSGRNPRL 126							
Db	65	HVKQPEEVTQAQLEGRSMNCPDYDKQTKTLFLFFTAVPGRVSEQHQLQIRVNVTRL 124							
Qy	127	CFICSDAGYSWSDVRDLTEEVIGPEVTHWATFVAGPGHGICQ--OSGRLLIIPAYAY-- 182							
Db	125	CRVSTSDYGMWSPVDLTETTTGTHODWATFVAGPGHCLQLRNAGSILLIPAYAYRL 184							
Qy	183	-----IPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGKLIKPMVTVVECEVAEVIKGAK 236							
Db	185	HPVHKPTPF-AFCF-----ISLDHGHTWELGNFVSEN-SLEQCAVEV-GTGA 228							
Qy	237	HPVLYCSARTPNRHRAEALSIDHGECFOKPVLSHOLCEPHGCOGSWSVF-CPLEIPGCG 295							
Db	229	HRVYLVNARSFIGARQQAQPNGLDFQDNQVYSKLVPEPPHCHGVSFAFHSPKDPCL 288							
Qy	296	QDLAGEDAPAIQGSPLLCSSVRPEAGTLESWLLYSHTNKKRVDLIGVYNOSPLEA 355							
Db	289	RHVAA-----YTHPTDSRNTNLGVYLNQTFPLDP 317							
Qy	356	ACWSRPWLHCGPGCYSDLA--ALENEG--LFGCLFECGTKECEQIAFLRTDREILSH 411							
Db	318	TANSEPTLLATGTCAVSDLIWGLGPDGSPFOGCLVESGNYDE--IIFLMFTLQAPPT 374							
Qy	412	VQG 414							
Db	375	VHG 377							

RESULT 4  
A54961  
sialidase, cytosolic - Chinese hamster  
C/Species: Crictetus griseus (Chinese hamster)  
C/Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 05-Nov-1999  
C/Accession: A54961  
R:Terrari, J.; Harris, R.; Warner, T. G.  
Glycobtology 4, 367-373, 1994  
A/Title: Cloning and expression of a soluble sialidase from Chinese hamster ovary cell  
A/Reference number: A54961; MUID:95036975  
A/Accession: A54961  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-379 <FER>



A:Cross-references: GB:U06143; NID:g509824; PIDN:AAA19746.1; PID:g509825

Query Match 29.5% Score 692; DB 2; Length 379;  
Best Local Similarity 38.68; Pred. No. 8.3e-52;  
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;  
Qy 13 LFOQEDKRGVYRIPALYVPPAHTFLAFAEKRSSKDEDALHLVRLRGTGQS--VQW 70  
Db 12 LFOQGD---YAIRPALIYLSKQTLAFAEKRLTKTDEHADLFVLRGSGYNADTHQVQW 68  
Qy 71 EPLKSLMKATLPGHRTMPCWPWERKSGYVYLFFICVOGHVTERQOIMSGRNPRLCEPIC 130  
Db 69 QAEEVVTQAYLEGRHSMSPCLDYKQRTFLFTFAVRGQISEHHQLQGTGVNVRCLCHIT 128  
Qy 131 SODAGYSWDRDLTEEVIGPEVTHWATFVAVGPGHIQLO--SGRLIIPAYAYIPIFWFF 188  
Db 129 STDHGKWSAVQDLDTDTTIGTHQDMATFVGGPGHCLQRLNRTASLLVPAYAYR----- 182  
Qy 189 CFRLP--YRARPMSLMYSDDLGATWHHGRILKPMVTVECEVAEYIGKAGHPVLVYCSART 246  
Db 183 --KOPPIHAPAPSAFCFLSHDHSWTWELGHEVS-QNSLECOVAEV-GTGAERVVYLNARS 238  
Qy 247 PNHRRAEALSDHGECEFOKPVLSHQLCPEPHGCGSVVSCFPLEIPGCGODLAGEDAPAI 306  
Db 239 CLGARVQAQSPNSGLDFODNQWVSKLVEPPKGCCHGSVIAF----- 278  
Qy 307 QOSPLCSSLVRPEPEAGTSLSESWLLYSHPTNKKRRVDLGIYLNQSPLEAACWSRPWILHC 366  
Db 279 -----PNMTSKADAL-DWMLITDTSDSRKNTGNYLVNKNPLDPTTWSAPTLLAT 328  
Qy 367 GPCGYSDLAAL-----ENEGFLGCFEFCGTCQCEQIAFRLFTDREILSHVOG 414  
Db 329 GICAYSDLQNHGHPDGSQPGCLYE-----SNNYEIVELMTLTKQAFPAVFG 377

RESULT 5  
T34999  
probable neuramidase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34999  
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: 221564  
A:Accession: T34959  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-648 <SEE>  
A:Cross-references: EMBL:AL034443; PIDN:CAA22361.1; GSPDB:GN00070; SCOEDB:SC4B5.07c  
C:Genetics:  
A:Gene: SCOEDB:SC4B5.07c

Query Match 11.0% Score 259; DB 2; Length 648;  
Best Local Similarity 25.3% Pred. No. 3e-14;  
Matches 128; Conservative 61; Mismatches 143; Indels 174; Gaps 30;  
Qy 5 TSCSFS-----SPLFOQE-----DKRG-VYTRIPALYVPPAHTFLAFA 42  
Db 15 TSCALAVSSPFAHAPGCDRAPAFGEQVLFDAARDPGGYACFRIPAVIRTTDGT-TLLAFA 73  
Qy 43 EKR-SSSKDEDAHLVLRGLRTGOSVOMEPLKSLMKATLPGHRTMPCWPWERKSGYVY 101  
Db 74 EGRVLDCAADGDDIDLVRSLDGGRT--WGPLRVVNDG--GGDTHGNPAPVYDRATGRVL 129  
Qy 102 LFFICVOGHVTERQOIMSGRNPRLCFT-C-----SODAGYSWSDVRDLTEVIGPE 152  
Db 130 LL-----ETYNAGRTSDACVPCARVPHVQHSDDGGRTWSAPRDLSPILPPD 178  
Qy 153 VTHWATFVAVGPGHGTQL-----OSGRLLIIPAYA-----YYIPFWFCFRLPYRARP- 198

Db 179 WNSW--YATGPVHGQVLTGGAHPGRLVGVVNAETWDSERSEMGPV-----PA 223  
Qy 199 -----HSLMIYSDDLGATWHHGRIL-----IKPMVTVECEVAEYIGRAGH 237  
Db 224 GGWRVVTANHAALVVDGGEHRTGATDTWPVAADGTFRQKPSLTLITERAD----- 276  
Qy 238 PVLVCSARTPN-----RIHRAEALSDHGECEFOKPVLSHQLCPEPHGCGSVVSCFPLEIPG 293  
Db 277 GALLVSGREENGTDPGHRTQALSDGGDSFAAPFA-----LP- 314  
Qy 294 GCODLAGEDAPAIQOSPLCSSLVRPEPEAGTSLSESWLLYSHPTNKKRRVDLGIYLNQSP 353  
Db 315 ---DLY---APQVQAVLRG-----NRLLSAPADPDRRTMTV--RSSRD 353  
Qy 354 EAACW-----SRPWILHCGPGYSDLAALNEGLFCGFCGTCQCEQIAF-RLF 403  
Db 354 GGATWDSADRGTVVTRDW-----AGYSDLVTVDDTV-GLLYEGGKTARDEIRFARLT 406  
Qy 404 TDREILSHVOG-DCSTFCM--NSEPS 426  
Db 407 ADR--LAPPRGPDPTTDLAANAAPA 430

RESULT 6  
A45244  
exo-alpha-sialidase (EC 3.2.1.18) - Micromonospora viridifaciens  
N:Alternate names: neuraminidase  
C:Species: Micromonospora viridifaciens  
C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 22-Oct-1999  
C:Accession: A45244  
R:Sakurada, K.; Ohta, T.; Hasegawa, M.  
J. Bacteriol. 174, 6896-6903, 1992  
A:Title: Cloning, expression, and characterization of the Micromonospora viridifaciens  
A:Reference number: A45244; MUID:93015752  
A:Accession: A45244  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-647 <SAK>  
A:Cross-references: GB:D01045; NID:g216782; PIDN:BAA00852.1; PID:d1001316; PID:g21678  
A:Note: sequence extracted from NCBI backbone (NCBIP:116820)  
C:Keywords: glycosidase; hydrolase

Query Match 8.6% Score 201.5; DB 2; Length 647;  
Best Local Similarity 24.0% Pred. No. 2.7e-09;  
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;  
Qy 12 PLFOQED-----KRGV-TYRIPALYVPPAHTFLAFAEKRSSKDEDALHLVLRRLGRT 64  
Db 49 PLYTEQDLAVNGREGFPNRYIPALT-VTPDGLLASVDGDRPTGIDAPGPNLSILQRRSTDG 107  
Qy 65 GQSVOMEPLKSLMKATLPGHRTMPCWPWERKSGYVYLFFICVOGHVTERQOIMSGRNP 124  
Db 108 GRTWGEQVWSAGQTTAPIKGFSDPYLVLDRETGTIFNF-----HVYSQGFAGSRPG 161  
Qy 125 -----RLCFTCSODAGYSWSDVRDLTEEVIGPEVTHWATFVAVGPGHIQLO-----S 171  
Db 162 TDPADPNVLHANVATSDGGLTWSH-RTITAD-ITPD-PGWRSRFAASGEGLRYGPHA 218  
Qy 172 GRL-----IIPAYAYIPIFWFCFRLPYRARPMSLMYISDDLGGATWHHGRILI-----KPMV 222  
Db 219 GRLIQOVTIINAAGAF-----QAVSVYSDDHGRTRWAGEAVGVGMDENK 262  
Qy 223 TVECEVAEYIGKAGHPVLVYCSARTPNHRRAEALSDHGECEFOKPVLSHQLCPEPHGCGS 282  
Db 263 TVELSDGRVLLNSRD-----SAR--SGYRKAVASVDGGHSGYGPVTIDRDLDPPTN--NAS 313  
Qy 283 VVSFCPLEIPGCGODLAGEDAPAIQOSPLCSSLVRPEPEAGTSLSESWLLYSHPTNKKRRV 342  
Db 314 IIRAFP-----DAP-----AGSARAKVLLFNSAASQTSR- 342  
Qy 343 DLGIYLNQSPLEAAC-----WSRPWILHCGPGCYSDLAALNEGLFCGFCGTCQCEQ 397

Db 343 -----SOGTIRMSCDDGQWTPVSKVQFGMSYSTLTALP-DGTGYLLYEPGT-----G 390  
QY 398 IAFRLFTDREILSHVQDC---STPCMNSEPKK 428  
Db 391 IRYANFN-----LAWLGICAPPTIPVALEPGQQ 420

RESULT 7  
T37097  
Probable secreted neuraminidase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T37097  
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.  
submitted to the EMBL Data Library, August 1999  
A:Reference number: Z21588  
A:Accession: T37097  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-748 <SAU>  
A:Cross-references: EMBL:AL109950; PIDN:CAB52948.1; GSPDB:GN00070; SCOEDB:SCJ4.14c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCJ4.14c

Query Match 8.4%; Score 197; DB 2; Length 748;  
Best Local Similarity 24.9%; Pred. No. 7.9e-09;  
Matches 103; Conservative 56; Mismatches 153; Indels 102; Gaps 21;

QY 22 VTYRIPALIVPPAHTFLFAEKR-SSSKDEDAHLVLRGLTGOSVQWEPKSLMKAT 80  
Db 76 VCFRAVAVKAADG-TLLAFAEGRICSCSDTASIDIVVKRYNGA----WSALQVVVARHS 130  
QY 81 LPHGRTMPCPVWERSKGVYLFPCVQGHVTERQOIMSGRNPALRCFICSQDAGYSWSD 140  
Db 131 -AGHIYHNVTVPVDAASGRVYLYTENYDIHR-----IASEDDGLHHTA 174  
QY 141 VRDLTEEVIGPEVTHWAT-----FAVPGCHGIQL---QSGRLIIPAYAYIPFWF 187  
Db 175 ADDISADV-----WSTAWGALYACQMATGPASALQLTHGRHAGRLVAG----- 217  
QY 188 FCFRLPYRAPHSL---MIYSDDLGATWHHGRLL---IKPMVTVECEVAEIVGKAGHPVLY 241  
Db 218 MTVRVAPGAAPANLGGALIVSDDGLTWRLGASSLGAEPVAGQA-ELS--LPERGDSLF 274  
QY 242 CSAR-----TPNHRRA-EALSDIHGECFQKP-VLSHQLCEPPHCCGCVSVVFCPLPIPGG 294  
Db 275 VTARNEEGSGTDRAYAVVSGDGLSFTSDFALLPMDLPCTGIGQASTIALREKNRDG- 333  
QY 295 CQDLAGEDAPAIQSPILCSSLVSRPEAGTLESWLLYSHPTNKKRRVDLGIYLNQSPLE 354  
Db 334 -----YDRALFAPVGNREDUTIRSSF-----DGGLTW-QDAAD 367  
QY 355 AACWSRPWILHCGPCGYSDLAALLENGLFCPLCEGCKQCEQIAFRFLTREI 408  
Db 368 GA-----LVKDGYSAYSSMTVL-GGTFGLYEAGTTKQYQDIRATTEADL 414

RESULT 8  
NMCLSS  
exo-alpha-sialidase (EC 3.2.1.18) precursor - Clostridium septicum  
N:Alternate names: neuraminidase  
C:Species: Clostridium septicum  
C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 18-Jun-1999  
C:Accession: S15994  
R:Rothe, B.; Rothe, B.; Roggentin, P.; Schauer, R.  
Mol. Gen. Genet. 226, 190-197, 1991  
A:Title: The sialidase gene from Clostridium septicum: cloning, sequencing, expression  
A:Reference number: S15994; MUID:91238693  
A:Accession: S15994  
A:Molecule type: DNA  
A:Residues: 1-1014 <MOL>

A:Cross-references: EMBL:X63266; NID:g40662; PIDN:CAA44916.1; PID:g40663  
C:Superfamily: Clostridium exo-alpha-sialidase  
C:Keywords: glycosidase; hydrolase  
F:1-36/Domain: signal sequence #status predicted <SIG>  
F:27-1014/Product: exo-alpha-sialidase #status predicted <MAT>

Query Match 7.9%; Score 186; DB 1; Length 1014;  
Best Local Similarity 27.5%; Pred. No. 1e-07;  
Matches 83; Conservative 37; Mismatches 102; Indels 80; Gaps 16;

QY 126 LCFICSQDAGYSWSDVRDLTEEVIGPEVTHWATF-AVGPCHGIQLQ-----SGRLIIPAYA 180  
Db 560 LSLIYSDDDGQWSDPIDUNKEV---KYDMRFLCTGPGKGHQIKTGRYAGRLFPVYL 615  
QY 181 YYIPFWFFCFRLPYRAPHSLMIYSDDLGATWH-----HGRLLI-----KPMVT--- 223  
Db 616 TNASGF-----QSSAVIYSDNGATWNIGETATDGRMDNGDRASAETITNTS 664  
QY 224 -----VECEVAEIVGKAGHPVLYCSARTPNRHRAP-ALSDIHGECFQKPVLSHQLCEPP 276  
Db 665 GGVGQLTECQVYVEM--PNCQLKMFMRNTGNSGRVRIATSFDCGATWEDDVVRDENIKER 722  
QY 277 HCGQGSVWSFCPLPTPGCGDLAGEDAPAIQSPILCSSLVSRPEAGTLESWLLYSHPT 336  
Db 723 Y-CQLSVINY-----SQKIDGKDA-----IIFAIPDAN-----YPNRV 754  
QY 337 NKKRRVLDL---GIYLNQSPLEAACWSRPWILHCGPCGYSDLAALLENGLFCGLFP-CGT 391  
Db 755 NQTVRVGLITENGSENGEPRYDIENRYKNVAPGTGYGSCISEMPN-GEIGLFYEGRS 813  
QY 392 KQ 393  
Db 814 RQ 815

RESULT 9  
I40866  
exo-alpha-sialidase (EC 3.2.1.18) - Clostridium perfringens  
C:Species: Clostridium perfringens  
C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Oct-1999  
C:Accession: I40866  
R:Traving, C.; Schauer, R.; Roggentin, P.  
Glycoconj. J. 11, 141-151, 1994  
A:Title: Gene structure of the 'large' sialidase isoenzyme from Clostridium perfringe  
A:Reference number: I40865; MUID:95102306  
A:Accession: I40866  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-694 <RES>  
A:Cross-references: EMBL:X87369; NID:g853805; PIDN:CAA60796.1; PID:g853807  
C:Genetics:  
A:Gene: nanH  
C:Keywords: glycosidase; hydrolase

Query Match 7.3%; Score 172; DB 2; Length 694;  
Best Local Similarity 20.8%; Pred. No. 1e-06;  
Matches 102; Conservative 65; Mismatches 143; Indels 180; Gaps 25;

QY 24 YRIPALIVPPAHTFLFAEKRSSSKD--EDALHLVLR---GLRT---GQSVQWEPK- 74  
Db 265 YRIPALFKTEGLTASIDARRHGGADAPNNDIDTAVRRSEGGKTWDEGQIIMDYPKS 324  
QY 75 SLMKATLPGHRTMPCPVWERSKGVYVLF-----IQCIV- 107  
Db 325 SVIDTTL-----IQDDETGRIFLLVTHPFPKYGFVWAGLGSFGFNIDGKEVLCY 374  
QY 108 -----QGHVTER-----QQIMSGRNPAP-----LCFI 129  
Db 375 DSSGKEFTYVENVYDKDGNKTEYTNALGDLFKNGTKIDNINSSAPLAKAGTSYNLV 434  
QY 130 CSQDAGYSWSDVRDLTEEVIGPEVTHWATF-AVGPCHGIQLQS-----GRLLIIPAYAYIP 184

Db 435 YSDDGKGTSEPNQINQV-----KKDMKFLGIAPGRGQIKNGECHKGRIVVPVY----- 485

QY 185 FWFECPRLPY---RAPHSLMIYSDDLGATW-----HHGRLIK-----PM 221

Db 486 -----YTNKEGKOSSAVIYSDSGKNWTIGESPNDNRKLENGKIINSKTLSDDAQ 536

QY 222 VTVECEVAEIVGAGHPVLVCSARTPNRHRAEALSIDHGECFOKPVLSHQLCRPPHGCQG 281

Db 537 LT-ECQVEM--PNGQLKLF--HRNLSGYLNIAISFDGGATWDETVEKQTNVLEPY--CQL 590

QY 282 SVVSFCPLEIPGCCQDLIAGEDAPAIQOSPLLCSSVRPEPEAGTLSWLLYSHPTNKKR- 340

Db 591 SVINY-----SQKVDGRDA-----VIFSNNARSRS 616

QY 341 ----RVDL----GIYLNQSPLEACHSRPWILHCGPCGYSDLAALNEGLFGCLFEGCTK 392

Db 617 NGTVRIGLINQVGYENGEPKYEFDKMYKLVKPYGYAYSCLTLSN--GNIGLLYEGTPS 675

QY 393 QECEQIAFRL 402

Db 676 EEMSYIEMNL 685

RESULT 10

A49227

sialidase - Actinomyces viscosus

C:Species: Actinomyces viscosus

C>Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 22-Oct-1999

C:Accession: A49227

R:Yeung, M.K.

A:Title: Complete nucleotide sequence of the Actinomyces viscosus T14V sialidase gene: A49227; MUID:93114861

A:Reference number: A49227

A:Contents: T14V

A:Accession: A49227

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-901 <YEU>

A:Cross-references: GB:L06898; NID:g289087; PIDN:AAA21932.1; PID:gl141852

A>Note: sequence extracted from NCBI backbone (NCBIN:121598, NCBIPI:121599)

Query Match 7.2%; Score 169; DB 2; Length 901;

Best Local Similarity 23.0%; Pred. No. 2.5e-06;

Matches 91; Conservative 50; Mismatches 155; Indels 100; Gaps 18;

QY 24 YRIPALIYVPPAHTFLAFABK-----RSSKDEDAHLVLRRLGRTGQSVQWPEPLKSLMK 78

Db 310 YRIPAITAPNGDLLISYDERPKDNGSGDAPNPNIHVQRRSTDGKKT--WSAPTVIHQ 367

QY 79 ATLPGHRT--MNPCEVWERSGVYVLFVFCVQGHVTERQ-----INSGRNPALCFI--- 129

Db 368 GTETGKKGVSDFSVYVDHQTGFNF-----HVKSYDQGWGSGRGTTDPENRGI 421

QY 130 --CSQDAGYSWSDVRLTEVIGPEVTH---WATFAVGPCHGIGLQ-----SGRLIIPAYA 180

Db 422 VSTSDNGTW-----THRTITADITKDKPWTARFAASGQGIQIHGPHAGRLV----- 470

QY 181 YIIPFWFECFRLPYRARP-----HSLMIYSDDLGATWHHGRLIKPMVTVECEVAEIVGKA 235

Db 471 -----QQYTIRTAGGAVQAVSVYSDHGHKTQWAG---TPICGTMDENKVV 516

QY 236 GHPVLVCSARTPNRHRAEALSIDHGECFOKPVLSHQLCRPPHGCQSVVSCFPLEIPGCC 295

Db 517 GSLMLNSRSDGSGFRKVAHSTGGQVWSEPVSDKNLPDSVDNAQ---IIRAFPNAA 571

QY 296 QDLAGEDAPAIQOSPLLCSSVRPEP-----EAGTLESWLLYSHPTNKKRRVDLGIYLNQSP 352

Db 572 -----DDPRAKVL--LLSHSPNRPWSDRGTTIS-----MISC 601

QY 353 LEAACHSRPWILHCGPCGYSDLAALNEGLFGCLFE 388

Db 602 DDCASWTTSKVPFHPFVGYYTI-AVQSDGSIGLLSE 636

RESULT 11

S20590

exo-alpha-sialidase (EC 3.2.1.18) - Actinomyces viscosus

C:Species: Actinomyces viscosus

C>Date: 22-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 22-Oct-1999

C:Accession: S20590

R:Henningsen, M.; Roggentin, P.; Schauer, R.

Biol. Chem. Hoppe-Seyler 372, 1065-1072, 1991

A:Title: Cloning, sequencing and expression of the sialidase gene from Actinomyces vi

A:Reference number: S20590; MUID:92162190

A:Accession: S20590

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-913 <HEI>

A:Cross-references: EMBL:X62276; NID:g39254; PIDN:CAA44166.1; PID:g39255

C:Keywords: glycosidase; hydrolase

Query Match 6.6%; Score 154; DB 2; Length 913;

Best Local Similarity 23.5%; Pred. No. 5.1e-05;

Matches 94; Conservative 48; Mismatches 150; Indels 108; Gaps 20;

QY 24 YRIPALIYVPPAHTFLAFABKRSKDEDA-----HLVLRRLGRTGQSVQWPEPLK 74

Db 310 YRIPA-IPPPMGT---CSSPTTSARTTATAAATPNPNIHVQRRSTDGKKT--HSAPT 363

QY 75 SLMKATLPGHRT--MNPCEVWERSGVYVLFVFCVQGHVTERQ-----INSGRNPALCF 128

Db 364 YIHQGTETGKKGVSDFSVYVDHQTGFNF-----HVKSYDQGWGSGRGTTDPENRGI 417

QY 129 I-----CSQDAGYSWSDVRLTEVIGPEVTH---WATFAVGPCHGIGLQ-----SGRLII 176

Db 418 IQAEVSTSDNGTW-----THRTITADITKDKPWTARFAASGQGIQIHGPHAGRLV- 470

QY 177 PAYAVYIPFWFECFRLPYRARP-----HSLMIYSDDLGATWHHGRLIKPMVTVECEVAE 231

Db 471 -----QQYTIRTAGGAVQAVSVYSDHGHKTQWAG---TPICGTMDENKVV 512

QY 232 IGKAGHPVLVCSARTPNRHRAEALSIDHGECFOKPVLSHQLCRPPHGCQSVVSCFPLEI 291

Db 513 ELSDGSLMLNSRSDGSGFRKVAHSTGGQVWSEPVSDKNLPDSVDNAQ---IIRAFPNAA 570

QY 292 PGCQDLAGEDAPAIQOSPLLCSSVRPEP---EAGTLESWLLYSHPTNKKRRVDLGIYL 348

Db 571 P-----DDPRAKVL--LLSHSPNRPWSDRGTTIS-----MISC 598

QY 349 NQSPLEAACHSRPWILHCGPCGYSDLAALNEGLFGCLFE 388

Db 599 -MSCDDGASWTTSKVPFHPFVGYYTI-AVQSDGSIGLLSE 636

RESULT 12

T30287

exo-alpha-sialidase (EC 3.2.1.18) - Streptococcus pneumoniae

N:Alternate names: neuraminidase

C:Species: Streptococcus pneumoniae

C>Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000

C:Accession: T30287

R:Camara, M.; Boulinois, G.J.; Andrew, P.W.; Mitchell, T.J.

Infect. Immun. 62, 3688-3695, 1994

A:Title: A neuraminidase from Streptococcus pneumoniae has the features of a surface

A:Reference number: Z20807; MUID:94341870

A:Accession: T30287

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1035 <CAM>

A:Cross-references: EMBL:X72967; NID:g587550; PID:g587553; PIDN:CAA51473.1

C:Genetics:

A:Gene: nana

C:Keywords: glycosidase; hydrolase



Db 203 FRIPGLVTNKGTLGVDYDVRNNSVDLQEHVDVGLSRSTGGKT--WEKMRPLAFGEF 260  
Qy 79 ATLPG--HRTNPNPCVWERKSGYVYLFICVQGHVTER-----QQINSGRNPALRCFICS 131  
Db 261 GGLPAGONGVGDPSILVDTKNNVYAAWTHGMGNORAMWSSHPGMDMNHNTAQLVLAKS 320  
Qy 132 ODAGYSWSDVR-DLTEEVIGPEVTHWATFAVPGHGIGLOQSGRLIIPAYAVYIPFWFFCF 190  
Db 321 TDDGKTSWGPPIINITEQVKDPS---WYFLQGPGRGITHSDGTLVFPQF----- 367  
Qy 191 RLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIGKAGHPVLVCSARTPNR- 249  
Db 368 -IDSTRVNPAGIMYSKDGKKNKKNHNYARTN-TTEAQVAEV-----EPGVLMNLNMRNRG 420  
Qy 250 -HRAEALSIDHGECFOKPVLSHOLCEPHGCGSVSFCPLIEPGCCODLAGEDAPAIQQ 308  
Db 421 GSRVAITKOLGKTWTEHESRK-----ALPE 447  
Qy 309 SPLLCSSVYRPEPEAGTLESWLLYSHPTNKKRRVLDLGIYLNQSPLEAACWS----- 359  
Db 448 SVCNASLISYKAKDNLGKDLLIESNPNTTKGRYNTIKISLD--GGVTWSPHQHLLDE 505  
Qy 360 -RPWILHCGPCGYSDLAALENEGLFGCLFE 388  
Db 506 GNNW-----GYSCLSMIDKETI-GILYE 527

RESULT 15  
A37234  
exo-alpha-sialidase (EC 3.2.1.18) - Clostridium sordellii  
C:Species: Clostridium sordellii  
C:Date: 30-Dec-1991 #sequence\_revision 30-Dec-1991 #text\_change 17-Mar-2000  
C:Accession: A37234  
R:Rothe, B.; Roggentin, P.; Frank, R.; Bloeker, H.; Schauer, R.  
J. Gen. Microbiol. 135, 3087-3096, 1989  
A:Title: Cloning, sequencing and expression of a sialidase gene from Clostridium sordellii  
A:Reference number: A37234; MUID:90132537  
A:Accession: A37234  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-404 <ROT>  
A:Cross-references: GB:M31584; NID:g144910; PIDN:AAA23280.1; PID:g144911  
C:Superfamily: trypanastigote-specific surface antigen  
C:Keywords: glycosidase; hydrolase

Query Match 6.2%; Score 145.5; DB 2; Length 404;  
Best Local Similarity 22.1%; Pred. No. 0.0001;  
Matches 80; Conservative 47; Mismatches 122; Indels 113; Gaps 16;

Qy 24 YRIPALIVPPAHTFLAPAEKR-SSSKDEADALHLVRLRGLTGQSVQWEPKLSMKATLP 82  
Db 54 FRIPSLQTLADG-TMLAFSDIRYGAEDHAYIDICAARKSTONGQTWDYKVTMENDRIDST 112  
Qy 83 GHRTMNPV-----WERKSGYVYLFVICVQGHVTERQOIMSGRNPALCFI 129  
Db 113 FSRVNDSTVYDTRGRIILIAGSNKKNWNA-----SSTSLRSDWSVOMV 158  
Qy 130 CSQDAGYSWSDVRDLT--EEVIGPEVTHWATFAVPGHGIGLOQSGRLIP-----A 178  
Db 159 YSDNGETWSDKVDLTNNKARIKQPSNTIGLAGVSGIVMSDGTIVMPIQIALRENNNA 218  
Qy 179 YAYIPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIGKAGHP 238  
Db 219 NNY-----SSVIYKDNGETWTMGNKVPDPKTSNNVIELDG----- 256  
Qy 239 VLYCSARTPNR-HRAEALSIDHGECFOKPVLSHOLCEPH-----CGQGSVVSF-- 286  
Db 257 ALIMSSRNDGKNRYASYISYDMGSTW-----EYDPLHNKISTGNSGCGQSGFIKVTA 309  
Qy 287 -----CPLEIPGG-----CQDLAGEDAPAIQQSPILCSSLVRPEAGT 324

Db 310 KDGHRGLGFISAPKNTKGGYVRDNITVYMIDPDDL-----SKGIRE-----LCS-----PYPEDGN 359  
Qy 325 LS 326  
Db 360 SS 361

Search completed: October 7, 2002, 14:59:51  
Job time: 73 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:59:58 ; Search time 17.45 seconds  
(without alignments)  
949.683 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTSFSPFLQEDRR.....LSHVQDCSTPGMNSPKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2348	100.0	428	NER3_BOVIN	Q7859 bos taurus
2	1916.5	81.6	428	NER3_HUMAN	Q9uq49 homo sapien
3	1518	64.7	418	NER3_RAT	Q99pw5 rattus norv
4	1501	63.9	418	NER3_MOUSE	Q9jmh7 mus musculu
5	711.5	30.3	380	NER2_HUMAN	Q9y3r4 homo sapien
6	705	30.0	379	NER2_MOUSE	Q9jmh3 mus musculu
7	705	30.0	379	NER2_RAT	Q84627 rattus norv
8	692	29.5	379	NER2_CRIGR	Q64393 cricetus
9	255.5	10.9	415	NER1_HUMAN	Q99519 homo sapien
10	248.5	10.6	409	NER1_RAT	Q99pw3 rattus norv
11	245.5	10.5	409	NER1_MOUSE	Q35657 mus musculu
12	201.5	8.6	647	NANH_MICVI	Q2834 micromosp
13	186	7.9	1014	NANH_CLOSE	P29767 clostridium
14	149	6.3	1035	NANA_STRPN	Q59959 streptococ
15	145.5	6.2	404	NANH_CLOSO	P15698 clostridium
16	128.5	5.5	266	NANH_BACFR	P31206 bacteroides
17	128	5.5	382	NANH_CLOPE	P10481 clostridium
18	107	4.6	1877	PKK5_MOUSE	Q04592 mus musculu
19	105	4.5	781	NANH_VIBCH	P37060 vibrio chol
20	99.5	4.2	2768	THYG_HUMAN	P01266 homo sapien
21	98	4.2	931	NRP2_MOUSE	Q35375 mus musculu
22	97	4.1	925	NRP2_RAT	Q35276 rattus norv
23	94.5	4.0	1216	AEGP_RAT	Q63191 rattus norv
24	94	4.0	381	NANH_SALTY	P29768 salmonella
25	94	4.0	2282	ZAN_RABIT	P57999 oryctolagus
26	91	3.9	697	NANB_STRPN	Q34727 streptococ
27	89.5	3.8	1255	PER2_HUMAN	Q15055 homo sapien
28	89	3.8	493	AMR2_DROEL	Q9njp0 drosophila
29	89	3.8	525	UL32_EBV	P03184 epstein-bar
30	87	3.7	493	1 AMYR_DROME	Q84808 drosophila
31	87	3.7	493	1 AMYR_DROTK	Q77018 drosophila
32	86.5	3.7	931	NRP2_HUMAN	Q60462 homo sapien
33	85	3.6	794	FURI_HUMAN	P09958 homo sapien

34	85	3.6	918	1 KPCM_MOUSE	Q62101 mus musculu
35	85	3.6	1025	1 CR2_MOUSE	P19070 mus musculu
36	84.5	3.6	731	1 SUFL_XENLA	Q9pun2 xenopus lae
37	84.5	3.6	1187	1 PTNE_HUMAN	Q15678 homo sapien
38	84.5	3.6	1255	1 ERB2_HUMAN	P04626 homo sapien
39	84	3.6	484	1 LEM2_PIG	P98110 sus scrofa
40	84	3.6	493	1 AMYR_DROAN	P18344 drosophila
41	84	3.6	2437	1 NOTC_BRARE	P46530 brachydanio
42	83.5	3.6	1141	1 SRE2_HUMAN	Q12772 homo sapien
43	83.5	3.6	1433	1 Y310_HUMAN	Q15027 homo sapien
44	83	3.5	493	1 AMYR_DROYA	Q76264 drosophila
45	83	3.5	494	1 AMYR_DROAP	Q77011 drosophila

#### ALIGNMENTS

RESULT 1  
NER3\_BOVIN  
ID NER3\_BOVIN STANDARD; PRT: 428 AA.  
AC Q97859;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)  
DE (N-acetyl-alpha-neuraminidase 3).  
GN NEU3.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=99143165; PubMed=9988745;  
RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,  
RA Sawada M.;  
RT "Molecular cloning and characterization of a plasma membrane-  
associated sialidase specific for gangliosides.";  
J. Biol. Chem. 274:5004-5011(1999).  
CC - FUNCTION: Plays a role in modulating the ganglioside content of  
the lipid bilayer at the level of membrane-bound sialyl  
glycoconjugates.  
CC - CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,  
alpha-(2->8)-glycosidic linkages of terminal sialic residues in  
oligosaccharides, glycoproteins, glycolipids, colominic acid and  
synthetic substrates.  
CC - SUBCELLULAR LOCATION: Membrane-associated (By similarity).  
CC - TISSUE SPECIFICITY: Expressed in brain.  
CC - SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC - SIMILARITY: CONTAINS 3 BNR REPEATS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
CC EMBL: AB008184; BAA75071.1; -.  
DR InterPro: IPR002860; BNR.  
DR Pfam: PF02012; BNR; 3...  
KW Hydrolyase; Glycosidase; Membrane; Repeat.  
REPEAT 129 140 BNR 1.  
FT REPEAT 203 214 BNR 2.  
FT REPEAT 254 265 BNR 3.  
FT SITE 24 27 FRIP MOTIF  
FT ACT\_SITE 25 25 BY SIMILARITY.  
FT ACT\_SITE 45 45 POTENTIAL.  
FT ACT\_SITE 50 50 POTENTIAL.  
FT ACT\_SITE 87 87 POTENTIAL.

```
FT ACT_SITE 225 225 POTENTIAL.
FT ACT_SITE 245 245 POTENTIAL.
FT ACT_SITE 341 341 BY SIMILARITY.
FT ACT_SITE 371 371 POTENTIAL.
FT ACT_SITE 388 388 POTENTIAL.
SQ SEQUENCE 428 AA; 47916 MW; 418B34F3245A8F21 CRC64;

Query Match 100.0%; Score 2348; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.5e-201;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEVTSFSFSPFQEDKRGVYRIPALIVVPPAHTFLAFAEKSSSKDEDALHLVLR 60
DB 1 MEEVTSFSFSPFQEDKRGVYRIPALIVVPPAHTFLAFAEKSSSKDEDALHLVLR 60
QY 61 GLRTGQSVQVEPLKSLMKATLPGHRTMNPVWPKSGYVYLFVFCVQGHVTERQQIMSG 120
DB 61 GLRTGQSVQVEPLKSLMKATLPGHRTMNPVWPKSGYVYLFVFCVQGHVTERQQIMSG 120
QY 121 RNPALRCFICSDAGYSWSDVRLTEVIGPETHWATFAVPGHGIQLOSGRLIIPAYA 180
DB 121 RNPALRCFICSDAGYSWSDVRLTEVIGPETHWATFAVPGHGIQLOSGRLIIPAYA 180
QY 181 YYIPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIKAGHPVL 240
DB 181 YYIPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIKAGHPVL 240
QY 241 YCSARTPNRHRAREALSIDHCECFQKPVLSHOLCEPHGCGSVVFCPLPIPGCCODLAG 300
DB 241 YCSARTPNRHRAREALSIDHCECFQKPVLSHOLCEPHGCGSVVFCPLPIPGCCODLAG 300
QY 301 EDAPAIQOSPILCSSVRPEPEAGTSLSESWLLYSHPTNKKRRVDLGIYLNOSPLEAACWSR 360
DB 301 EDAPAIQOSPILCSSVRPEPEAGTSLSESWLLYSHPTNKKRRVDLGIYLNOSPLEAACWSR 360
QY 361 PWILHCGPGYSIDLALENGLFCGCEGTCQKCEQIAFRFTDRILSHVQDCSTPG 420
DB 361 PWILHCGPGYSIDLALENGLFCGCEGTCQKCEQIAFRFTDRILSHVQDCSTPG 420
QY 421 MNSEPSKK 428
DB 421 MNSEPSKK 428

RESULT 2
NER3_HUMAN
ID NER3_HUMAN . STANDARD; PRT; 428 AA.
AC Q9UQ49; Q9NOE1;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99335353; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside
RT sialidase.";
RL Biochem. Biophys. Res. Commun. 261:21-27(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti E., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B.,
RA Croci G., Preti A., Hallabio A., Tettamanti G., Borsani G.;
```



Db	301	ADAPTQSS--PGSSLKEEEAGTPSSWLLYSHTSRKQKVDGIYLNQNPLEAACWSR	359
Qy	361	PWLHCPCGCGYSDLAALLENELGFLCFCGCTKQCEQTAFRLFTDREILSHVQGDCTSPG	420
Db	360	PWLHCPCGCGYSDLAALLENELGFLCFCGCTKQCEQIAFLFTTHREILSHLQGDCTSPG	419
Qy	421	MNSEPSK	427
Db	420	RN--PSQ	424
RESULT 3			
NER3_RAT			
ID	NER3_RAT	STANDARD;	PRT; 418 AA.
AC	Q9PW5;		
DT	01-MAR-2002	(Rel. 41, Created)	
DT	01-MAR-2002	(Rel. 41, Last sequence update)	
DE	01-MAR-2002	(Rel. 41, Last annotation update)	
DE	Slialidase 3 (EC 3.2.1.18)	(Membrane slialidase)	(Ganglioside slialidase)
DE	(N-acetyl-alpha-neuraminidase 3).		
GN	NEU3.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
ON	NCBI_TaxID=10116;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;		
RX	MDLINE=J0192669; PubMed=1162581;		
RA	Hasegawa T., Feijoo Carnero C., Wada T., Itoyama Y., Miyagi T.;		
RT	"Differential expression of three slialidase genes in rat		
RT	development.";		
RL	Biochem. Biophys. Res. Commun. 280:726-732(2001).		
CC	-1- FUNCTION: Plays a role in modulating the ganglioside content of		
CC	the lipid bilayer at the level of membrane-bound sialyl		
CC	glycoconjugates (By similarity).		
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,		
CC	alpha-(2->8)-glycosidic linkages of terminal sialic residues in		
CC	oligosaccharides, glycoproteins, glycolipids, colominic acid and		
CC	synthetic substrates.		
CC	-1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).		
CC	-1- TISSUE SPECIFICITY: Expressed in brain, cardiac muscle and weakly		
CC	in liver.		
CC	-1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.		
CC	-1- SIMILARITY: CONTAINS 3 BNR REPEATS.		
-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
-----			
CC	EMBL; AB026841; BAB32440.1; -.		
DR	InterPro; IPR002860; BNR.		
DR	Pfam; PF02012; BNR: 3		
KW	Hydrolase; Glycosidase; Membrane; Repeat.		
FT	REPEAT 129 140	BNR 1.	
FT	REPEAT 201 212	BNR 2.	
FT	REPEAT 252 263	BNR 3.	
FT	SITE 24 27	FRIP MOTIF.	
FT	ACT_SITE 25 25	BY SIMILARITY.	
FT	ACT_SITE 45 45	POTENTIAL.	
FT	ACT_SITE 50 50	POTENTIAL.	
FT	ACT_SITE 87 87	POTENTIAL.	
FT	ACT_SITE 223 223	POTENTIAL.	
FT	ACT_SITE 243 243	POTENTIAL.	
FT	ACT_SITE 339 339	BY SIMILARITY.	
FT	ACT_SITE 369 369	POTENTIAL.	
FT	ACT_SITE 386 386	POTENTIAL.	
SQ	SEQUENCE 418 AA; 446980 MW; 7CC46F2E5952E240	CRC64;	

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC EMBL; AB026842; BAA92868.1; -  
DR MGD; MGI:1355305; Neu3.  
DR InterPro: IPR002860; BNR.  
DR Pfam; PF02012; BNR; 3.  
KW Hydrolase; Glycosidase; Membrane; Repeat.  
FT REPEAT 129 140 BNR 1.  
FT REPEAT 203 214 BNR 2.  
FT REPEAT 254 265 BNR 3.  
FT SITE 24 27 FRIP MOTIF.  
FT ACT\_SITE 25 25 BY SIMILARITY.  
FT ACT\_SITE 45 45 POTENTIAL.  
FT ACT\_SITE 50 50 POTENTIAL.  
FT ACT\_SITE 87 87 POTENTIAL.  
FT ACT\_SITE 223 223 POTENTIAL.  
FT ACT\_SITE 243 243 POTENTIAL.  
FT ACT\_SITE 339 339 BY SIMILARITY.  
FT ACT\_SITE 369 369 POTENTIAL.  
FT ACT\_SITE 386 386 POTENTIAL.  
SQ SEQUENCE 418 AA; 46846 MW; 64853FC963FE7686 CRC64;

Query Match 63.9%; Score 1501; DB 1; Length 418;  
Best Local Similarity 67.3%; Pred. No. 8.6e-126;  
Matches 282; Conservative 47; Mismatches 86; Indels 4; Gaps 2;

QY 1 MEEVTSFSSPLFOQEDKRGVYRIPALIVYPPAHTFLAFAEKRSKDEDAHLVLR 60  
DB 1 MEEVPPYSLSSTLFQEQSGVYRIPALDYLPTHTFLAFAEKRTSVRDEDAACLVLR 60  
QY 61 GLRTGQSVQWPELKSMLKATLPGHRTMPCPVWERKSGVYVFFICVQGHVTERQOIMSG 120  
DB 61 GLMKRSVQWGPORLLMEATLPGHRTMPCPVWEKNTGRVYLFICVGRHVTERCQIVWG 120  
QY 121 RNPARLCFCSQDAGYSWSDVRLTEVIGPEVTHWATFVAVGPHGIGLQSGRLIIPAYA 180  
DB 121 KNAARLCFLCSEDAGSGWGEVKDLTEVIGSEVKRWATFVAVGPHGIGLQSGRLIIPAYA 180  
QY 181 YIIPFFCFRLPYRAPHSLMIYSDDLGATWHHGRILKIPMTVVECEVAEVIKAGHPVL 240  
DB 181 YVYSRWFCLFAC--SVKPHSLMIYSDDFGVTHHGKFIQVPTGECQVAEVAETAGNPVL 238  
QY 241 YCSARTPNRHRRAEALSIDHGECFQKPVLSHOLCEPPHGCQSGVSVFPCLEIPGCGQDLAG 300  
DB 239 YCSARTPSRFAEAFSDSGCFQKPTLNPQLHEPTGCGVSVFRLPKMPTYODSIG 298  
QY 301 EDAPAIQSPLLCSSVRPEPEAGTLESWLLYSHPTNKKRRVDGLYLNQSPLEAACWSR 360  
DB 299 KGAPATQKCPLLDSPLEVEKGAETPSATWLLYSHPTSKRRINLGIYNNRNPLEVNCWSR 358  
QY 361 PWILHCGPCGYDLAALENGLFCGECCTKQECQIAFRLETDRILLSHVQGDSTP 419  
DB 359 PWILNRGPGSYDLAAVEEDLVACLFECGERNEYERIDFCLESDHEVLVS--CEDCTSP 415

RESULT 5  
NER2\_HUMAN STANDARD; PRT; 380 AA.  
AC Q9Y3R4;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Sialidase 2 (EC 3.2.1.18) (Cytosolic sialidase) (N-acetyl-alpha-neuraminidase 2).  
GN NEU2.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99208670; PubMed=10191093;  
RA Monti E., Preti A., Rossi E., Ballabio A., Borsani G.;  
RT "Cloning and characterization of NEU2, a human gene homologous to  
RL rodent soluble sialidases.";  
RL Genomics 57:137-143(1999).  
CC -!- FUNCTION: Hydrolyzes sialylated compounds.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,  
CC alpha-(2-8)-glycosidic linkages of terminal sialic residues in  
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and  
CC synthetic substrates.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, fetal liver and  
CC embryonic carcinoma cell line NT2B1.  
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -!- SIMILARITY: CONTAINS 2 BNR REPEATS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Y16535; CAB41449.1; -  
DR MIM; 605528; -  
DR InterPro; IPR002860; BNR.  
DR Pfam; PF02012; BNR; 2.  
KW Hydrolase; Glycosidase; Repeat.  
FT REPEAT 127 138 BNR 1.  
FT SITE 197 208 BNR 2.  
FT ACT\_SITE 20 23 FRIP MOTIF.  
FT ACT\_SITE 21 21 BY SIMILARITY.  
FT ACT\_SITE 46 46 POTENTIAL.  
FT ACT\_SITE 243 243 POTENTIAL.  
FT ACT\_SITE 304 304 BY SIMILARITY.  
FT ACT\_SITE 334 334 POTENTIAL.  
FT ACT\_SITE 355 355 POTENTIAL.  
SQ SEQUENCE 380 AA; 42230 MW; 9D18F1041A2D4F44 CRC64;

Query Match 30.3%; Score 711.5; DB 1; Length 380;  
Best Local Similarity 40.7%; Pred. No. 9.6e-56;  
Matches 170; Conservative 59; Mismatches 114; Indels 75; Gaps 16;

QY 10 SSPLFOQED--KRGV--TYRIPALIVYPPAHTFLAFAEKRSKDEDAHLVLRGG--LR 63  
DB 3 SLPLVOKESVFGCAHAYRIPALLYLPQGSLLAFARASKDEHAELIVLRGVDAP 62  
QY 64 TGSQVQWEPLKSLMKTLPGLHRTMPCPVWERKSGVYVFFICVQGHVTERQOIMSGRNP 123  
DB 63 THQ-VQWQAEVVAQARLDGHRSMNCPPLYDAQTGTLFFIAIPGVTEQQQLQTRNV 121  
QY 124 ARLCFTCSQDAGYSWSDVRLTEVIGPEVTHWATFVAVGPHGIGIQL--QSGRLIIPAYAY 181  
DB 122 TRLCQVTSIDHGTSWSPRLDTAAIGPAYREWSSTFVAVGPHGICLNDARSLLWPAYAY 181  
QY 182 Y-----IPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRILKIPMTVVECEVAEVI 233  
DB 182 RKLHPTQRIIPS-AFCF-----LSHDGRTWARGHFV-AQDTLEQCAVEV-- 224  
QY 234 KAGHP-VLYCSARTPNRHRRAEALSIDHGECFQKPVLSHQLCE-PPHGCQSGVSVFPCLEI 291  
DB 225 ETGEQRVVTILNARSHLRARVQAQSTNDGLDFQESQLVKVLVEPPPGCGQSGVSVF---- 279  
QY 292 PGCCQDLAGEDAPAIQOQSPLLCSSVRPEPEAGTLES-ULLYSHPTNKKRRVDGLYLNQ 350  
DB 280 -----PSRSGPGSPAQWLLYTHPTHSWRADLGLAYLP 313

```

QY 351 SPLEACNAGSRPWILHCGPCGYSDLAAL-----ENEGLFCLFEFGTKQBCEGIAFLFT 404
|   || | | | | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
DB 314 RPPAPEAWSEPVLAKGCAYSDLQSMGTGPDGSPFLGCLYEAE---NDYEEIVLFMT 368

RESULT_6
NER2_MOUSE
ID NER2_MOUSE STANDARD; PRT; 379 AA.
AC Q9JMH3; O99NA3;
CT 01-WAR-2002 (Rel. 41, Created)
DT 01-WAR-2002 (Rel. 41, Last sequence update)
DT 01-WAR-2002 (Rel. 41, Last annotation update)
DE Sialidase 2 (EC 3.2.1.18) (Cytosolic sialidase) (N-acetyl-alpha-neuraminidase 2) (MURINE thymic sialidase) (MTS) (Mouse skeletal muscle sialidase) (MSS).
DE NEU2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
SEQUENCE FROM N.A.
R1 TISSUE=Brain;
RC MEDLINE=99262165; PubMed=10329453;
RX Pronda C.L., Zeng G., Gao L., Yu R.K.;
RA "Molecular cloning and expression of mouse brain sialidase.";
RL Biochem. Biophys. Res. Commun. 258:727-731(1999).
RN [2]
SEQUENCE FROM N.A.
R1 TISSUE=Skeletal tissue;
RC MEDLINE=20179907; PubMed=10713120;
RX Hasegawa T., Yamaguchi K., Wada T., Takeda A., Itoyama Y., Miyagi T.
RA "Molecular cloning of mouse ganglioside sialidase and its increased expression in Neuroza cell differentiation.";
RL J. Biol. Chem. 275:8007-8015(2000).
RN [3]
SEQUENCE FROM N.A.
R1 TISSUE=Thymus;
RC MEDLINE=21391602; PubMed=11500029;
RX Kotani K., Kuroiwa A., Salto T., Matsuda Y., Koda T.,
RA Kitajimoto-Ochiai S.;
RT "Cloning, chromosomal mapping, and characteristic 5'-UTR sequence o murine cytosolic sialidase.";
RL Biochem. Biophys. Res. Commun. 286:250-258(2001).
CC -1- FUNCTION: Hydrolyzes sialylated compounds.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-, alpha-(2->8)-glycosidic linkages of terminal sialic residues in oligosaccharides, glycoproteins, glycolipids, colominic acid and synthetic substrates.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: Highly expressed in heart.
CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -1- SIMILARITY: CONTAINS 2 BNR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabor between the Swiss Institute of Bioinformatics and the EMBL outst the European Bioinformatics Institute. There are no restrictions use by non-profit institutions as long as its content is in modified and this statement is not removed. Usage by and for com entities requires a license agreement (See http://www.isb-sib.ch/an or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF139059; ; NOT_ANNOTATED_CDS.
DR EMBL; AB028023; BAA92867.1; -.
DR EMBL; AB048604; BAB39152.1; ALT_INIT.
DR MGD; MGI:134417; Neu2.
DR InterPro; IPRO02860; BNR.
DR Pfam; PF02012; BNR; 2.
KW Hydrolase; Glycosidase; Repeat.
FT REPEAT 127 138 BNR 1.
FT REPEAT 197 208 BNR 2.
FT SITE 20 23 FRIP MOTIF.
FT SITE 20 23

```

RT cytosolic sialidase gene.";  
RL Glycobiology 5:511-516(1995).  
CC -!- FUNCTION: Hydrolyzes sialylated compounds.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,  
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in  
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and  
CC synthetic substrates.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -!- SIMILARITY: CONTAINS 2 BNR REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D16300; BAA03805.1; -;  
DR EMBL; D50606; BAA09169.1; -;  
DR InterPro: IPR002860; BNR.  
DR Pfam; PF02012; BNR; 2.  
KW Hydrolase; Glycosidase; Repeat.  
FT REPEAT 127 138 BNR 1.  
FT SITE 197 208 BNR 2.  
FT SITE 20 23 FRIP MOTIF.  
FT ACT\_SITE 21 21 BY SIMILARITY.  
FT ACT\_SITE 46 46 POTENTIAL.  
FT ACT\_SITE 243 243 POTENTIAL.  
FT ACT\_SITE 303 303 BY SIMILARITY.  
FT ACT\_SITE 333 333 POTENTIAL.  
FT ACT\_SITE 354 354 POTENTIAL.  
SQ SEQUENCE 379 AA; 42382 MW; 55583C7043CA9784 CRC64;

Query Match 30.0%; Score 705; DB 1; Length 379;  
Best Local Similarity 39.5%; Pred. No. 3.6e-55;  
Matches 167; Conservative 60; Mismatches 126; Indels 70; Gaps 13;

QY 12 PLFQOE---DKRGVTRIPALIVVPPAHTFLAFAPKRSKDEDAHLVLRGLRTGQS- 67  
DB 5 PVLQKETLFHTEYVAYRIPALLYLKKQKTLAPAPKRSRTDEHAEILVLRGSGYNGATN 64  
QY 68 -VQWELKSLMKATLPGHRTMPCPVWKRSGVYVLFVFCVQGHVTERQOIMSGRNPRL 126  
DB 65 HVKMQPEEVVQALQEGHSMNCPCLYDKQTKTLFFIAVPCRVSEHQLOLQTRVNVTRL 124  
QY 127 CFTCSODAGYSWSDRLTEEVIGPEVTHWATFVAGPGHGIOL--OSGRLLIIPAYAY-- 182  
DB 125 CRVTSYDGMNNSPVQDLTETTTIGSTHQDWATFVAGPGHCLQLRNAGSLLPAYAYRKL 184  
QY 183 -----IPWFVFCFLPYRAPHSLMIYSDLLGATWHHGRLLKPMVTVEVAEVIKAG 236  
DB 185 HPVHKPTFP-ATCF-----ISLDHGHWTWELGNFVSEN-SLEQCAEV-GTGA 228  
QY 237 HPVLYCSARTPNRHRAEALSIDHGEQFQVLSHOLCEPPHGCQGSWSF-CPLEIPGCC 295  
DB 229 HRVVLNARSFTGARVQAOASPDGLDFQDNQVVKLVEPPHCHGSGVAFHSPTSKPDCL 288  
QY 296 QDLAGEDAPAIQOOSPLLCSSVRPEPAGTSLSESWLYSHPTNKKRVDIGIYLNQSPLEA 355  
DB 289 RHVAA-----YTHPTDSRNTNLGVNLQTPLDP 317  
QY 356 ACSWRPWLHCPCGYSDLA--ALENEG--LFGCLFECTKQCEBOIAFRLTDRILSH 411  
DB 318 TAWSEPTLATGTCAYSDLOIWLGLDGPSPQFGCLYESNYDE---IIFLMTLQKQAPT 374  
QY 412 VQG 414  
DB 375 VHG 377

RESULT 8  
NER2\_CRIGR  
ID NER2\_CRIGR STANDARD; PRT; 379 AA.  
AC Q64393;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DE 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Sialidase 2 (EC 3.2.1.18) (Cytosolic sialidase) (N-acetyl-alpha-  
DE neuraminidase 2).  
GN NEU2.  
OS Cricetulus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetulus.  
OX NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC CCL61;  
RX MEDLINE=95036975; PubMed=7949662;  
RA Ferrari J., Harris R., Warner T.G.;  
RT "Cloning and expression of a soluble sialidase from Chinese hamster  
RT ovary cells: sequence alignment similarities to bacterial  
RT sialidases";  
RL Glycobiology 4:367-373(1994).  
CC -!- FUNCTION: Hydrolyzes sialylated compounds.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,  
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in  
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and  
CC synthetic substrates.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -!- SIMILARITY: CONTAINS 2 BNR REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U06143; AAA19746.1; -;  
DR InterPro: IPR002860; BNR.  
DR Pfam; PF02012; BNR; 2.  
KW Hydrolase; Glycosidase; Repeat.  
FT REPEAT 127 138 BNR 1.  
FT REPEAT 197 208 BNR 2.  
FT SITE 20 23 FRIP MOTIF.  
FT ACT\_SITE 21 21 BY SIMILARITY.  
FT ACT\_SITE 46 46 POTENTIAL.  
FT ACT\_SITE 243 243 POTENTIAL.  
FT ACT\_SITE 303 303 BY SIMILARITY.  
FT ACT\_SITE 333 333 POTENTIAL.  
FT ACT\_SITE 354 354 POTENTIAL.  
SQ SEQUENCE 379 AA; 41962 MW; B5AFFBC6B6BE88B1 CRC64;

Query Match 29.5%; Score 692; DB 1; Length 379;  
Best Local Similarity 38.6%; Pred. No. 5.1e-54;  
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;

QY 13 LFQOEDKRGVTRIPALIVVPPAHTFLAFAPKRSKDEDAHLVLRGLRTGQS--VQW 70  
DB 12 LFQTDG---YAYRIPALIVLSKQKTLAPAPKRLTKTDEHADFLVLRGSGYNADTHQVQW 68  
QY 71 EPLKSLMKATLPGHRTMPCPVWKRSGVYVLFVFCVQGHVTERQOIMSGRNPRLCFTC 130  
DB 69 QAEVYVQAYLEGHRSMSPCPLYDKQTKTLFFIAVRCQIASEHHQLOLQTVGVNTRLCHIT 128  
QY 131 SODAGYSWSDRLTEEVIGPEVTHWATFVAGPGHGIOLQ--SGRLIIPAYAYIIPFWFF 188  
DB 129 STDHCKTWSAVQDLDTTIGSTHQDWATFVAGPGHCLQLRNAGSLLPAYAYR----- 182









CC -|- DOMAIN: A C-terminal internalization signal (YGR1) appears to  
CC allow the targeting of plasma membrane proteins to endosomes.  
CC -|- PPM: N-glycosylated (Probable).  
CC -|- PPM: Phosphorylation of tyrosine within the internalization signal  
CC results in inhibition of sialidase internalization and blockage on  
CC the plasma membrane.  
CC -|- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -|- SIMILARITY: CONTAINS 4 BNR REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; Y11412; CAA72215.1; -;  
DR EMBL; U93702; AAC53536.1; -;  
DR EMBL; AF109906; AAC84167.1; -;  
DR EMBL; BC004566; AAH04566.1; -;  
DR HSP; Q02834; LEUT.  
DR MGD; MGI:97305; Neul.  
DR InterPro; IPR002860; BNR.  
DR Pfam; PF02012; BNR; 4.  
KW Hydrolase; Glycosidase; Signal; Repeat; Glycoprotein.  
FT SIGNAL 1 41  
FT CHAIN 42 409  
FT REPEAT 106 117  
FT REPEAT 166 177  
FT REPEAT 225 236  
FT REPEAT 241 352  
FT SITE 71 74 FRIP MOTIF.  
FT SITE 406 409 INTERNALIZATION SIGNAL.  
FT ACT\_SITE 72 72 BY SIMILARITY.  
FT ACT\_SITE 91 91 POTENTIAL.  
FT ACT\_SITE 129 129 POTENTIAL.  
FT ACT\_SITE 258 258 POTENTIAL.  
FT ACT\_SITE 274 274 POTENTIAL.  
FT ACT\_SITE 335 335 BY SIMILARITY.  
FT ACT\_SITE 364 364 POTENTIAL.  
FT ACT\_SITE 388 388 POTENTIAL.  
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 209 209 L -> I (REDUCED ACTIVITY).  
FT CONFLICT 107 107 R -> K (IN REF. 2).  
FT CONFLICT 113 113 S -> C (IN REF. 2).  
FT CONFLICT 117 121 STAFI -> PTGFM (IN REF. 2).  
FT CONFLICT 172 173 GI -> AF (IN REF. 2).  
FT CONFLICT 344 344 F -> L (IN REF. 2).  
FT CONFLICT 351 351 Q -> L (IN REF. 4).  
SQ SEQUENCE 409 AA; 44591 MW; 416BFD5BE27B8893 CRC64;

Query Match 10.5%; Score 245.5; DB 1; Length 409;  
Best Local Similarity 27.1%; Pred. No. 2.5e-14;  
Matches 112; Conservative 46; Mismatches 136; Indels 119; Gaps 22;  
QY 23 TYRIPALVPPAHTFLAFAE-KRSSCKDEDAHLVLRGLRTGQSVQWEPKLSMKATL 81  
D 70 TFRIP-LITATPRCTLFAFAEARKKSASDEGAKFIAMRS--TDQGTWSTAFIVDDGE 126  
QY 82 PGHTMPCPWERKSGSYVLYFF-ICVQGHVTERQQIMSGRNPARLCFFICSDAGYSWSD 140  
D 127 ASDGLNLGAVVNDVTGIVELIVTLC--AHKVNQC-----VASTMLVMSKDGISWSP 177  
QY 141 VRDLTEEVIGPEVTHWATFAVGCHGILQ-----SGRLIIPAYAYIPFFWFCERLPYRA 196  
D 178 PRLNSVD-IGTEM-----FAPGSGIGKQREPCKGRLIVCGHGHTLDRGVFC----- 224

QY 197 RPHSLMIYSDDLGLATWHHGRLLIKPMVTVECEVAEIVGKAGHPVLYCSARTPNRRAEALS 256  
D 225 -----LLSDDHGHASWHYG-----TGVSIGP----- 244  
QY 257 IDHGECEQKPKVSLHQLCEPHGCGSVVPCPLPIPGCCQDLAGED-----APAIQOS 309  
D 245 -----EQQPKHDHF--NPDECO-----PYELPGSVIINARNQNNHYHCRIVLRS 289  
QY 310 PLLCSSVRP-----EPE-----AGTL--SESWLLYSHPTNKKRRVLDGLYLNQSPLEA 355  
D 290 YDADCTLRPRDVTDPDLVDPVVAAGALATSSGIVFFSNPAHPFRVNL--TLKWSFSG 347  
QY 356 ACWSRPWI-LHCPCGYSDLAALENE-----GLFGCLFECGCTKQCEQIA 399  
D 348 TSWQKERVQWMPGPGSGYSSLTALENSTDGKKQPPOLF-VLYEKLNLRYTESIS 399  
RESULT 12  
ID NANH\_MICVI STANDARD; PRT; 647 AA.  
AC Q02834;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sialidase precursor (EC 3.2.1.18) (Neuraminidase).  
GN NEBA.  
OS Micromonospora viridifaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Micromonosporineae; Micromonosporaceae;  
OC Micromonospora.  
OX NCBI\_TaxID=1881;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-647.  
RC STRAIN=ATCC 31146;  
RX MEDLINE=93015752; PubMed=1400240;  
RT "Cloning, expression, and characterization of the Micromonospora  
RT viridifaciens neuraminidase gene in Streptomyces lividans.";  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
RC STRAIN=ATCC 31146;  
RX MEDLINE=96164436; PubMed=8591030;  
RA Gaskell A., Crenell S., Taylor G.;  
RT "The three domains of a bacterial sialidase: a beta-propeller, an  
RT immunoglobulin module and a galactose-binding jelly-roll.";  
RL Structure 3:1197-1205(1995)  
CC -|- FUNCTION: TO RELEASE SIALIC ACIDS FOR USE AS CARBON AND  
CC ENERGY SOURCES FOR THIS NON-PATHOGENIC BACTERIUM WHILE IN  
CC PATHOGENIC MICROORGANISMS, SIALIDASES HAVE BEEN SUGGESTED TO  
CC BE PATHOGENIC FACTORS.  
CC -|- CATALYTIC ACTIVITY: HYDROLYSIS OF 2,3-, 2,6- AND 2,8-GLYCOSIDIC  
CC LINKAGES JOINING TERMINAL NON-REDUCING N- OR O-ACYLNEURAMINYL  
CC RESIDUES TO GALACTOSE, N-ACETYLHEXOSAMINE, OR N- OR O-ACETYL  
CC NEURAMINYL RESIDUES IN OLIGOSACCHARIDES, GLYCOPROTEINS,  
CC GLYCOLIPIDS OR COLOMINIC ACID.  
CC -|- SUBCELLULAR LOCATION: Secreted.  
CC -|- INDUCTION: BY N-ACETYLNEURAMINIC ACID, COLOMINIC ACID, AND SIALIC  
CC ACID.  
CC -|- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -|- SIMILARITY: CONTAINS 7 BNR REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D01045; BAA00852.1; -;  
DR FIR; A45244; A45244.  
DR PDB; 1EUR; 11-JAN-97.



```
DR PDB; 1EUS; 11-JAN-97.
DR PDB; 1EUT; 11-JAN-97.
DR PDB; 1EUV; 11-JAN-97.
DR InterPro; IPR002860; BNR.
DR InterPro; IPR000421; FA58_C.
DR Pfam; PF02012; BNR; 5.
DR SMART; SM00231; FA58C; 1.
KW Hydrolase; Glycosidase; Signal; Repeat; 3D-structure.
FT SIGNAL 1 37
FT CHAIN 38 647
FT ACT_SITE 260 260
FT ACT_SITE 370 370
FT REPEAT 102 113
FT REPEAT 175 186
FT REPEAT 239 250
FT REPEAT 287 298
FT REPEAT 348 359
SQ SEQUENCE 647 AA; 68830 MW; DCC1FE5BC935B8AD CRC64;

Query Match 8.6%; Score 201.5; DB 1; Length 647;
Best Local Similarity 24.0%; Pred. No. 3.4e-10;
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;

Qy 12 PLFOQED-----KRGV-TYRIPALIVPPAHTFLFAFAKRRSSKDEDAHLHLVL-RRGLRT 64
Db 49 PLYTEQDLAVNGREGFPNVRIPALT-VTPDGLLASVDGRPTGIDAPGPNLSILQRSTDG 107
Qy 65 QGSQVQWELKSLMKATLPGHRTMPCPWRKSGVYVLFCTVCGHVTQRQIMSGRNP 124
Db 108 GRTWGEQVWSAGQTTAPIKGFSYLVDRGTGIFNF-----HVYSQRQGFAGSRPG 161
Qy 125 -----RLCFICSQDAGYSWSDVRDLTERVIGPEVTHWATFVAGCHGIGLOQ----S 171
Db 162 TDPADPNVLHANVATSTGGLTWSH-RTITAD-ITPD-PGWRSRFAAAGEGQLRYGPHA 218
Qy 172 GRL-----IIPAYAYIIPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLI-----KPMV 222
Db 219 GRLIQOYTIINAAGAF-----QAVSVYSDDHGRTRAGEAVGCMDENK 262
Qy 223 TVECEVAEVIKAGHPVLYCSARTPNRHRAEALSDHGECKFKPVLSHOLCEPPHCGCS 282
Db 263 TVELSDGRVLLNSRD-----SAR--SGYRKVAVSTDGGHSYGPVTIDRLDPDTN--NAS 313
Qy 283 VVSFCPLEIPGCGQDLAGEDAPAIQOQSPILCCSSVRPEPAGTLESWLLYSHTNKKRRV 342
Db 314 IIRAPF-----DAP-----AGSARAKVLLFSNAASQTSR- 342
Qy 343 DLGIYLNOSPLEAAC-----WSRPWILHCGCGYSDLAALENEGLFCIFGCGTKQCEEQ 397
Db 343 -----SQGTIRMSCDDGQTPVSKVPFGMSYSTLTALP-DGYTGLLYEPCT-----G 390
Qy 398 IAFRLFTREILSHVQDC---STPGMNSPESKK 428
Db 391 IRYANFN-----LAWLGGICAPFTIPDVALEPGQQ 420

RESULT 13
NANH_CLOSE
AC NANH_CLOSE STANDARD; PRT; 1014 AA.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sialidase precursor (EC 3.2.1.18) (Neuraminidase).
OS Clostridium septicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1504;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC 0034714;
RX MEDLINE=91238693; PubMed=2034213;
```

```
RA Rothe B., Rothe B., Roggentin P., Schauer R.;
RT "The sialidase gene from Clostridium septicum: cloning, sequencing,
RT expression in Escherichia coli and identification of conserved
RT sequences in sialidases and other proteins.";
CC Mol. Gen. Genet. 226:190-197(1991).
CC -|- FUNCTION: SIALIDASES HAVE BEEN SUGGESTED TO BE PATHOGENIC FACTORS
CC IN MICROBIAL INFECTIONS.
CC -|- CATALYTIC ACTIVITY: HYDROLYSIS OF 2,3-, 2,6- AND 2,8-GLYCOSIDIC
CC LINKAGES JOINING TERMINAL NON-REDUCING N- OR O-ACYLNEURAMINYL
CC RESIDUES TO GALACTOSE, N-ACETYLHEXOSAMINE, OR N- OR O-ACETYLATED
CC NEURAMINYL RESIDUES IN OLIGOSACCHARIDES, GLYCOPROTEINS,
CC GLYCOLIPIDS OR COLOMINIC ACID.
CC -|- SUBCELLULAR LOCATION: Periplasmic.
CC -|- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -|- SIMILARITY: CONTAINS 4 BNR REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63266; CAA44916.1; -.
DR PIR; S15994; NMCLSS.
DR HSP; O02834; 1EUR.
DR InterPro; IPR002860; BNR.
DR InterPro; IPR000421; FA58_C.
DR InterPro; IPR004124; sialidase_N.
DR Pfam; PF02012; BNR; 4.
DR Pfam; PF02973; sialidase_N; 1.
KW Hydrolase; Glycosidase; Periplasmic; Signal; Repeat.
FT SIGNAL 1 26
FT CHAIN 27 1014
FT REPEAT 431 442
FT REPEAT 563 574
FT REPEAT 627 638
FT REPEAT 700 711
FT REPEAT 1014 AA; 110652 MW; C4F49233473A2FAD CRC64;
SQ SEQUENCE 1014 AA; 110652 MW; C4F49233473A2FAD CRC64;

Query Match 7.9%; Score 186; DB 1; Length 1014;
Best Local Similarity 27.5%; Pred. No. 1.4e-08;
Matches 83; Conservative 37; Mismatches 102; Indels 80; Gaps 16;

Qy 126 LCFICSQDAGYSWSDVRDLTERVIGPEVTHWATF-AVGGHGIQIQ-----SORLIIPAYA 180
Db 560 LSLIYSDDDGQTSDDPIDLNKEV---KTDWMRFGLTGPKGHGIKTRYAGRLFPVYL 615
Qy 181 YYIPWFECFRLPYRAPHSLMIYSDDLGATWH-----HGRLI-----KPMVT--- 223
Db 616 TNASGF-----OSSAVIYSDNGATWNIGETATDGRMDNGDRASAEITITNTS 664
Qy 224 -----VECEVAEVIKAGHPVLYCSARTPNRHRAE-ALSDHGECKFKPVLSHOLCEPP 276
Db 665 GGVGQLTECQVVM--PNGQLKMFNRNTGGNSGRVRIATSPDGGATWEDDVVRDINREP 722
Qy 277 HCGQSVVSFCPLEIPGCGQDLAGEDAPAIQOQSPILCCSSVRPEPAGTLESWLLYSHT 336
Db 723 Y-CQLSVINY-----SQKIDGKDA-----IIFAIPDAN-----YPNRV 754
Qy 337 NKKREVDL-----GIYLNOSPLEAACWSRPWILHCGCGYSDLAALENEGLFCFLPE-CGT 391
Db 755 NGTVRVGLITENGSYENGEPRYDIENRYKNVAVPTGYTGYSCUSEMPN-GETGLFEGRGS 813
Qy 392 KQ 393
Db 814 RQ 815

RESULT 14
NANA_STRPN
```

ID NANA\_STRPN STANDARD; PRT; 1035 AA.  
AC Q59959; Q54722;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sialidase A precursor (EC 3.2.1.18) (Neuraminidase A).  
GN NANA.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R36A / NCTC 10319;  
RX MEDLINE=94341870; PubMed=8063384;  
RA Camara M., Boulnois G.J., Andrew P.W., Mitchell T.J.;  
RT "A neuraminidase from Streptococcus pneumoniae has the features of a  
surface protein.";  
RL Infect. Immun. 62:3688-3695(1994).  
RN [2]  
RP SEQUENCE OF 882-1035 FROM N.A.  
RC STRAIN=SEROTYPE 6;  
RX MEDLINE=96326329; PubMed=8759848;  
RA Berry A.M., Lock R.A., Paton J.C.;  
RT "Cloning and characterization of nanB, a second Streptococcus  
pneumoniae neuraminidase gene, and purification of the NanB enzyme  
from recombinant Escherichia coli.";  
RL J. Bacteriol. 178:4854-4860(1996).  
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL  
NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.  
CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -1- SIMILARITY: CONTAINS 4 BNR REPEATS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X72967; CAA51473.1; -  
DR EMBL; U43526; AAC44391.1; -  
DR HSP; 002834; LEUR.  
DR InterPro; IPR002860; BNR.  
DR InterPro; IPR001791; Laminin\_G.  
DR Pfam; PF02012; BNR; 4.  
DR Pfam; PF02973; sialidase\_N; 1.  
DR SMART; SM00282; LamG; 1.  
DR Hydrolase; Glycosidase; Cell wall; Repeat; Signal.  
FT SIGNAL 1 ? POTENTIAL.  
FT CHAIN ? 1035 SIALIDASE A.  
SQ SEQUENCE 1035 AA; 114741 MW; C5B8A2D7A1E212F3 CRC64;  
  
Query Match 6.3%; Score 149; DB 1; Length 1035;  
Best Local Similarity 19.8%; Pred. No. 2.7e-05;  
Matches 103; Conservativity 63; Mismatches 159; Indels 200; Gaps 22;  
  
QY 9 FSSPLFOQEDKRGV-TYRIPALIVPPAHTFLAFAEKRSKDEDAHLVLR-----GLR 63  
DB 330 FESGRCKPKNGDKSVRIAPALLKTDGTLIAGADERLHSSDWGIDGMVIRSENGKT 389  
QY 64 TGSQVQNEPLKLMKATLPG-----HRTMPCPVWE----- 94  
DB 390 WGRVTTNLRDNPKASDPSIGSVNIDMVLVDPEPKRIFSIYDMPEKGIFGMSQK 449  
QY 95 -----RKSGYVYLFFICVOGHVTEROOINSGRNP----- 124  
DB 450 E5AYKKIDGTYQILYREGKGYATTRENGTVY-----TPDGKATDYRVVDPVKPYSK 505

QY 125 -----RLCFICSQDAGYSWSVDRLDTEEVIGPEVTH 155  
DB 506 GDLYKGNOLLNIGYFTTNKTSPFRIAKDSYLMMSYDDGKTWSAPQDITPMV-----KAD 561  
QY 156 WATF-AVPGHGIGLOQS-----GRLLIPAY-----AYIPFFFCFLPYRAPHSLMIYSD 206  
DB 562 WMKFLGVGPGTGIIVLRNGPHKGRILIPVYTTNNVSHL-----NGSQSRIIYSD 610  
QY 207 DLGATWHHGLIKPMVTVECEVAEIVGKAGHPVLYCSARTPNRRHAE-----ALSIDHGE 261  
DB 511 DHGKTHWAGEAVN-----DNROVDGQKTH-----SSTMNRRRAONTSTVVLNNGD 657  
QY 262 CFQKPVLSHOLCEPPHCGCGSVWSFCPLTPGGCQDLAGEDA-----PAIOOSP 310  
DB 658 -----VKLFMRGLTGLDLQVATSKDGGVTWEKDKIKRYPOVKDVI 695  
QY 311 LLCSSVRPEEAGTLESWLLYSHPTNKKR-----RVDLG-----IYLNQSPLEAACWS 359  
DB 696 VQMSAIHTMHE---GKEYITLSNAGGPKRENGVMVHLARVEENGELTWLKHNPQK---- 747  
QY 360 RPWILHCGPGCGYSLAALANEGLFGCLFECGTK-OECEQIAFLRF 403  
DB 748 -----GEFAYNSLOELGN-GEYGILYEHTKQCNAYTISFRKF 784  
  
RESULT 15  
NANH\_CLOSO STANDARD; PRT; 404 AA.  
ID NANH\_CLOSO  
AC P15698;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sialidase precursor (EC 3.2.1.18) (Neuraminidase).  
OS Clostridium sordellii.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1505;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-45.  
RC STRAIN=GL2;  
RX MEDLINE=90132537; PubMed=2693593;  
RA Rothe B., Roggert P., Frank R., Bloeker H., Schauer R.;  
RT "Cloning, sequencing and expression of a sialidase gene from  
Clostridium sordellii GL2.";  
RL J. Gen. Microbiol. 135:3087-3096(1989).  
CC -1- FUNCTION: SIALIDASES HAVE BEEN SUGGESTED TO BE PATHOGENIC FACTORS  
IN MICROBIAL INFECTIONS.  
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL  
NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.  
CC -1- SUBCELLULAR LOCATION: Periplasmic.  
CC -1- PTM: IT IS POSSIBLE THAT THE SIALIDASE IS CLEAVED IN FRONT OF A  
CYSTEINE WITHIN THE LEADER PEPTIDE, FORMING A GLYCERIDE  
THIOETHER BOND WHICH LINKS THE PROTEIN TO THE MEMBRANE. A SECOND  
PROTEOLYTIC CLEAVAGE RELEASES THE MATURE EXTRACELLULAR PROTEIN.  
CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -1- SIMILARITY: CONTAINS 5 BNR REPEATS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M31584; AAA23280.1; -  
DR PIR; A37234; A37234.  
DR HSP; P29768; 2SIL.  
DR InterPro; IPR002860; BNR.  
DR Pfam; PF02012; BNR; 5.  
DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; UNKNOWN\_1.  
KW Hydrolase; Glycosidase; Periplasmic; Signal; Repeat.  
FT SIGNAL 1 27





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:59:18 ; Search time 50.13 Seconds

(without alignments)  
1476.996 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTSFSFSLFQEDKR.....LSHVQGDCTPGMNSFSSKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	259	11.0	648	2 Q9ZBW1	Q9Zbw1 streptomyc
2	197	8.4	748	2 Q9SIV9	Q9Siv9 streptomyc
3	172	7.3	694	2 Q59310	Q59310 clostridium
4	169	7.2	901	2 Q44562	Q44562 actinomyc
5	154.5	6.6	544	2 Q45145	Q45145 bacteroides
6	154	6.6	913	2 Q59164	Q59164 actinomyc
7	150.5	6.4	404	2 P94675	P94675 clostridium
8	142.5	6.1	404	2 P94674	P94674 clostridium
9	140.5	6.0	404	2 P94676	P94676 clostridium
10	137.5	5.9	1003	2 Q9AHP2	Q9Ahp2 arcanobacte
11	132.5	5.6	762	5 Q27701	Q27701 macrobactera
12	125	5.3	740	16 Q97Q99	Q97q99 streptococc
13	123	5.2	568	12 Q911I6	Q911i6 frog adenov
14	120	5.1	382	2 Q59311	Q59311 clostridium
15	117.5	5.0	773	2 P77848	P77848 clostridium
16	115.5	4.9	643	5 Q9BHJ5	Q9Bhj5 trypanosoma

17	113.5	4.8	642	5 Q26969	Q26969 trypanosoma
18	113.5	4.8	771	5 Q9GSF0	Q9gsf0 trypanosoma
19	111.5	4.7	642	5 Q26968	Q26968 trypanosoma
20	109.5	4.7	642	5 Q26967	Q26967 trypanosoma
21	109.5	4.7	771	5 Q9GU83	Q9gu83 trypanosoma
22	109.5	4.7	964	5 Q26963	Q26963 trypanosoma
23	109.5	4.7	1070	2 Q9EZV7	Q9ezv7 pasteurella
24	108.5	4.6	642	5 Q26966	Q26966 trypanosoma
25	108	4.6	1170	2 Q9AJR8	Q9ajr8 erysipelothe
26	107.5	4.6	517	12 Q9YU51	Q9yus1 turkey aden
27	106.5	4.5	879	5 Q00773	Q00773 trypanosoma
28	106.5	4.5	1060	5 Q26964	Q26964 trypanosoma
29	103	4.4	556	4 Q9UMP5	Q9ump5 homo sapien
30	102	4.3	680	11 Q55001	Q55001 mus musculu
31	101.5	4.3	700	5 Q08672	Q08672 trypanosoma
32	100	4.3	567	4 Q9BWE0	Q9bwe0 homo sapien
33	100	4.3	567	4 Q9BUZ6	Q9buz6 homo sapien
34	95.5	4.1	389	10 Q23780	Q23780 chlamydomon
35	95.5	4.1	558	12 Q84169	Q84169 oliveros vi
36	94.5	4.0	2820	5 Q9VLT6	Q9vit6 drosophila
37	92	3.9	567	4 Q9NZH2	Q9nzh2 homo sapien
38	92	3.9	628	4 Q9BTF0	Q9btf0 homo sapien
39	92	3.9	841	11 P97484	P97484 mus musculu
40	92	3.9	1080	16 Q9CM43	Q9cm43 pasteurella
41	91.5	3.9	700	5 Q26965	Q26965 trypanosoma
42	91	3.9	1299	12 Q9YUJ1	Q9ytj1 ateline her
43	90	3.8	381	12 Q91GK2	Q91gk2 epiphyas po
44	90	3.8	1436	10 Q9LFY6	Q9lfy6 arabidopsis
45	90	3.8	1572	5 Q44938	Q44938 haemochus

#### ALIGNMENTS

RESULT 1

Q9ZBW1 ID Q9ZBW1 PRELIMINARY; PRT; 648 AA.

AC Q9ZBW1; 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE PUTATIVE NEURAMIDASE.

GN SC4B5.07C.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.

OX NCBI\_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Seeger K.J., Harris D.;

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J., Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

DR EMBL; AL034443; CAA22361.1; -.

DR HSSP; Q02834; 1EUR.

DR InterPro; IPR002860; BNR.

DR InterPro; IPR001791; Laminin\_G.

DR Pfam; PF02012; BNR; 5.

DR SMART; SM00282; LamG; 1.

SEQUENCE 648 AA; 68216 MW; F1CD5835276679EB CRC64;

```
Query Match 11.0%; Score 259; DB 2; Length 648;
Best Local Similarity 25.3%; Pred. No. 1.4e-16;
Matches 128; Conservative 61; Mismatches 143; Indels 174; Gaps 30;

Qy 5 TSCSFS-----SPLFOE-----DKRG-VYIRIPALIVPPAHTFLAFA 42
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 15 TSCALAVSPPFAHARPGRAPAGEQVLEDAARDPGYACFRIPAIYRTTGD-TLLAFA 73
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 43 EKR-SSSKDEDALHLVLRRTGQSVQWEPKLSMKATLPCRMTNCPVWERKSGVYV 101
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 74 EGRVLDCAADGDIDIVLRSLDGGRT--WGPLRVVNDG--GGDTHGNPAPVVDATGRVL 129
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 102 LFFICVQGHVTRQOIMSGRNPRLCFI-C-----SQDAGYSWSDVRLITEEVIGPE 152
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 130 LL-----ETYNAGRTSDSACVPCARVPHVOHSDPGRTWSAPRLSPILPPD 178
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 153 VTHWATFVAGPGHGIOL-----QSGRLIIPAYA-----YIPFWFCFRLPYRARP- 198
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 179 WNSW--YATGPVHGVLGGAHPGRLVGVNAETWDGERSEMGVP-----PA 223
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 199 -----HSLMIYSDDLGATWHHGR-----IKPMVTVECEVAEIVIGKAGH 237
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 224 GGWRVTANHAALVVSDDGGEHWRGTATWPVAADGTFROKPSLTLTERAD----- 276
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 238 PVLYCSARTPN----RHRAEALSIDHGECFOKPVLSHOLCEPPHGCQSGSVVFCPLETPG 293
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 277 GALLVSGREENGTDPGHTQALSRDGGDSFAAPFA-----LP- 314
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 294 GQDLAGEDAPAIQSPILCSSVRPEPEAGTLSWLLYSHPTNKKRVDLGIYLNQSP 353
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 315 ---DLV---AFQVQAVLURL-----NRLLSAPADPRRTMTV--RSSRD 353
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 354 EAACW-----SRPWILHCGPGYSDLAALNEGLFCGCTKQCEQIAF-RLF 403
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 354 GGATWDSADRGTVYTRDW-----AGSYDLVYDDTV-GLLYEGGRTDARDEIRFARLT 406
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 404 TDREILSHVQG-DCSTPGM--NSEPS 426
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 407 ADR--LAPPRGPDPPTPDLAANAAPA 430
   |||: : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 2

```
Q9S1V9 PRELIMINARY; PRT; 748 AA.
AC Q9S1V9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE SECRETED NEURAMINIDASE.
GN SCJ4-14C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C.; Harris D.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D.; Parkhill J.; Barrell B.G.; Rajandream M.A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M.; Kleser H.M.; Denapante D.; Eichner A.; Cullum J.;
RA Kinashi H.; Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
```

```
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. Microbiol. 21:77-96(1996).
DR EMBL; AL109950; CAB52948.1; -.
DR HSSP; Q02834; 1EUT.
DR InterPro; IPR002860; BNR.
DR Pfam; PF02012; BNR; 4.
SQ SEQUENCE 748 AA; 79060 MW; 20E829A5AB867571 CRC64;

Query Match 8.4%; Score 197; DB 2; Length 748;
Best Local Similarity 24.9%; Pred. No. 1.8e-10;
Matches 103; Conservative 56; Mismatches 153; Indels 102; Gaps 21;

Qy 22 VTYRIPALIVPPAHTFLAFAEKR-SSSKDEDALHLVLRRLTGTQSVQWEPKLSMKAT 80
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 76 VCFRAPAVVKAADG-TLIAFAEGRIGSCDSTASIDIVVKRYVNGA----WSALQVVARHS 130
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 81 LPGAHTMPCPVWRKSGVYVLFVICVQGHVTERQOIMSGRNPRLCFICQDAGYSWSD 140
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 131 -AGHIYHNVTVPVDAASGRVVVLYTENYDHIHR-----IASEDDGLHHTA 174
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 141 VRDLTEEVIGPEVTHWAT-----FVGPFGHGIQL-----QSGRLIIPAYYIPFWF 187
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 175 ADDISADV-----WSTAWGALYAGQMATGPASAIQLTHGRHAGRLVAG----- 217
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 188 FCFRLPYRAPHSL---MIYSDDLGATWHHGR-----IKPMVTVECEVAEIVIGKAGHPVLY 241
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 218 MTVRVAPCAAPANLGGALIYSDGGLTWRLGASSILGAEPVGAQ-ELS--LPERGDGSLF 274
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 242 CSAR-----TPNRHRA-EALSIDHGECFOKP-VLSHQLCEPPHGCQSGVVSFCLPEIPGG 294
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 275 VTARNEEGSDTRAVYAVSGDQGLSFTSDFALLPMDLEGTGTAQSTLALREKNRDG- 333
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 295 CODLAGEDAPAIQSPILCSSVRPEPEAGTLSWLLYSHPTNKKRVDLGIYLNOSPLE 354
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 334 -----YDRALFAAPVGNREDLTISSF-----DGGLTW-QDAAD 367
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 355 AACRSRWILHCGPGYSDLAALNEGLFCGCTKQCEQIAFRTTQREI 408
   |||: : : : : : : : : : : : : : : : : : : : : : : :
Db 368 GA-----LVKDGYSAYSSMTVL-GGDTFGILYEAGTKQYQDIRFATFTADL 414
   |||: : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 3

```
Q59310 PRELIMINARY; PRT; 694 AA.
AC Q59310;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EXO-ALPHA-SIALIDASE PRECURSOR (EC 3.2.1.18).
GN NANH.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A99;
RX MEDLINE=95102306; PubMed=7804004;
RA Traving C.; Schauer R.; Roggentin P.;
RT "Gene structure of the 'large' sialidase isoenzyme from Clostridium
RT perfringens A99 and its relationship with other clostridial nanH
RT proteins.";
RL Glycoconj. J. 11:141-151(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A99;
RA Schauer R.;
RT SEQUENCE FROM N.A.
RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; X87369; CAA60796.1; -.
DR HSSP; Q02834; 1EUT.
DR InterPro; IPR002860; BNR.
DR InterPro; IPR004124; sialidase_N.
```







```

Qy 24 YRIPALIVYPPAHTFLAFAEKR--SSSKDEDALHLVLRRLGTQGSQVQWEPLKSLMKATLP 82
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 54 FRIPSLQTLADG-TMLAFSDIRYNGAEAHAYIDIGAAKSTDNGQTDWYKTVMMENDRIDST 112
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Qy 83 GHRTMNPcv-----WERKSGYVYLLFFTCVQGHVTERQOIIMSGRNPRLCFI 129
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 113 FSRVMDSTTVVTDTRILLAGSWNKGNA-----SSTTSLRSDWSQVMV 158
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Qy 130 CSODAGYSWSDVRDLT--EEVIGPEVTHMATFAVGPCHGILQSGRLIIP-----A 178
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 159 YSDNGETWSDKVDLTNTNKARIKQNSNTTGWLAGVSGSIVMSDGTIVMPIQIALRENN 218
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Qy 179 YAYIPEWFFCFRLPYRARPHSLMIYSDDLGATWHHGRILKPMVTVECEVAEYIGRAGHP 238
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 219 NNYI-----SSVIKSKDIGETWPMGNKWMDPKTSENVIELDG----P 257
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Qy 239 VLCSARTPNHRAREALSIOHGCFOKPVLSHOLCEPPH-----GCQGSV 284
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 258 LIMSRRNDGKNYRASYISYDLGSTW-----EYVDPLHNKISTGNGSGCQGSFI 305
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|

RESULT 9
P94676
ID P94676 PRELIMINARY; PRT; 404 AA.
AC P94676;

```

```

RESULT 10
Q9AHP2
ID Q9AHP2 PRELIMINARY; PRT; 1003 AA.
AC Q9AHP2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEURAMINIDASE NANH (EC 3.2.1.18).
GN NANH.
OS Arcanobacterium pyogenes.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Actinomycineae; Actinomycetaceae; Arcanobacterium.
OX NCBI_TaxID=1661;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21295094; Pubmed=11401983;
RA Jost B.H., Songer J.G., Billington S.J.;
RT "Cloning, expression, and characterization of a neuraminidase gene
RT from Arcanobacterium pyogenes.";
RL Infect. Immun. 69:4430-4437(2001).
DR EMBL; AF298154; AAK15462.1; -.
DR HSSP; Q02834; LEUR.
DR InterPro; IPR002860; BNR.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF02012; BNR; 5.

```

DR	EMBL; U46666; AAC47263.1; -.
DR	InterPro: IPR002860; BNR.
DR	InterPro: IPR001791; Laminin_G.
DR	InterPro: IPR004124; sialidase_N.
DR	Pfam: PF02012; BNR: 4.
DR	Pfam: PF02973; sialidase_N; 1.
DR	SMART; SM0282; LamG; 1.
KW	Hydrolase; Glycosidase; Signal; Repeat.
FT	SIGNAL 1 28 POTENTIAL.
FT	CHAIN 29 762 SIALIDASE_L.
FT	REPEAT 328 339 BNR MOTIF.
FT	REPEAT 511 522 BNR MOTIF.
FT	REPEAT 571 582 BNR MOTIF.
FT	REPEAT 620 631 BNR MOTIF.
FT	ACT_SITE 293 293 BY SIMILARITY.
FT	ACT_SITE 611 611 BY SIMILARITY.
FT	ACT_SITE 673 673 BY SIMILARITY.
SEQUENCE	762 AA; 82982 MH; CQC547C7A8632B37 CRC64;
Query Match 5.6%; Score 132.5; DB 5; Length 762;	
Best Local Similarity 24.5%; Pred. No. 0.00035;	
Matches 67; Conservative 40; Mismatches 94; Indels 73; Gaps	
QY	131 SQDAGYSWSVDRLTEVIGPEVTHNATFAVGHGHIQL-----QSGRLLIPAYAIYPFW 186           :       :         :         :         :
Db	513 SDDEGASMDL-DIVSS-FKPEYSKF--LVVGPGLQISTGENAGRLLVPLYS----- 562           :       :         :         :         :
QY	187 FFCFRFLPYRRPHSLMTIYSDDLGATWHHGRLL--IKPNVTVECEVAEICKAGHPVLVCSA 244 :: :: :: :: ::   :   :   :   :   :   :   :   :   :   :   :
Db	563 -----KSSAEGLFMYSDDHGNNTYVEADNLTGATAEAQIVEMPGS-----LKTYL 610 
QY	245 RTPNRHRAELSIDHCGECKFKPVLSHQLCBPFGCGSVSFPCLEIPGCGQDLAGEDAP 304     :             :           :       :       :       :       :
Db	611 RTGSNCIAEVTSIDGETSWDRPLOGISTSTGTQLSVINY-----SQPIDCK--P 660 
QY	305 AIQQSPLLCCSVRPEPEAGLTSESWLLYSHPTN--KKRRVDLGIY-----LNQSPLEA 355           :       :       :       :       :       :       :       :
Db	661 AI-----ILSSPNA-----TNGRKNGKIWLGVNDTGTGTGDIKYSVE- 697           :       :       :       :       :       :       :       :
QY	356 ACMSRPWILHCGPCGYSDLAALE-NEGLFCCLFE 388   :   :           :       :       :       :       :       :       :
Db	698 --WKYSYAVDTPQMGSYSYSCLAELPDGQVGLLYE 729   :   :           :       :       :       :       :       :       :
RESULT	12
Q97Q99	PRELIMINARY; PRT; 740 AA.
ID	Q97Q99
AC	Q97Q99;
DT	01-OCT-2001 (TrEMBLrel. 18, Created)
DT	01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	NEURAMINIDASE, PUTATIVE.
DN	SP1326.
OS	Streptococcus pneumoniae.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OX	Streptococcus.
NCBI_TaxID=	1313;
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=TIGR4;
RX	MEDLINE=21357209; PubMed=11463916;
RA	Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA	Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA	Burkin A.S., Winn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA	Mayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA	Holtzaple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA	McDonald L.A., Feldbush T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT	"Complete genome sequence of a virulent isolate of Streptococcus
RL	pneumoniae.";
RL	Science 293;498-506(2001).

```
DR EMBL; AB007431; AAK75424.1; -.
DR TIGR; SP1326; -.
DR InterPro; IPR002860; BNR.
DR InterPro; IPR003975; Shal_channel.
DR InterPro; IPR004124; sialidase_N.
DR Pfam; PF02012; BNR; 4.
DR Pfam; PF02973; sialidase_N; 1.
DR PRINTS; PR01497; SHALCHANNEL.
KW Complete proteome.
SQ SEQUENCE 740 AA; 82437 MW; 51101837A55FA4A7 CRC64;

Query Match
Best Local Similarity 5.3%; Score 125; DB 16; Length 740;
Matches 46; Conservative 25; Mismatches 69; Indels 26; Gaps 7;

QY 126 LFCICSDAGYNSVDRLTEEVIGPEVTHWATFVPGHG-IQLQSGRLIIPAYAYIP 184
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 503 IAYISSNDHGWS-APTLLPTMG--LNRNAPY-LGPGRIESTGRILPSY----- 553
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 185 FWFECFRLPYRAPHSLMIYSDDLGLATWHHGRLLIKPMV-TVECEVAEVIKAGHPVLYCS 243
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 554 -----TKESAFIYSDNGASVKVVPPLPSSWSAEAQFVEL-----SPGVIQAY 598
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 ARTPNRHRAEALSIDHGECFOKPVLSHQLCPEPHGCGGVSVCPL 289
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 599 MRTNCKIAVLTSKDAGTWSAPEYKLFVSNPSYGTQLSIINYSOL 644
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
ID Q9II16 PRELIMINARY; PRT; 568 AA.
AC Q9II16;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DR EMBL; AF224336; AAF86921.1; -.
DR InterPro; IPR002860; BNR.
DE SIALIDASE.
OS Frog adenovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Siadenovirus.
OX NCBI_TaxID=114102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20451109; PubMed=10993931;
RA Davison A.J., Wright K.M., Harrach B.;
RT "DNA sequence of frog adenovirus.";
RL J. Gen. Virol. 81:2431-2439(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Davison A.J., Wright K.M., Harrach B.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF224336; AAF86921.1; -.
DR InterPro; IPR002860; BNR.
DR Pfam; PF02012; BNR; 5.
SQ SEQUENCE 568 AA; 62542 MW; 64279B9837985FAD CRC64;

Query Match
Best Local Similarity 5.2%; Score 123; DB 12; Length 568;
Matches 70; Conservative 51; Mismatches 137; Indels 46; Gaps 13;

QY 8 SFSSP-----LFOQDKRGVYRIPALIVPPAHTFLAEKRSSKQEDALH---LVLR 59
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 SYTVPARWTVFSGDLGTFYRIPELICL-KSGVLIAGGDACYNAFDD--FRTCTIAVA 120
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 RGLRTGSQVOMEPLKSLMATLPGHRTMNPVWKRSGYVYLFFICVGHVTVTQOIMS 119
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 RSEDDGGCTGDKOCPVPAFAVFKARFLDACVV-EDTLGRVHLFV-----YEENQHIS 174
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 GRNPA-RLCFCISQDAGYNSVDRLTEEVIGPEVTHWATFVPGHG-IQLQSGRLIIPA 178
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 175 STDPNWDVFHVHSLDEKSSPSSSLKSLAKDSERYFQS-----PGNGIVMKDGTVVVPC 230
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 YAYIIPFWFFCFRLPYRAPHSLMIYSDDLGLATWHHGRLLIKPMVTVCEVAEVIKAGHP 238
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
DB 231 QAW-----LKTTSGIWTSFYSKDNQGVYTWTRCATDLPVNSSENMAE-YPAGQL 281
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 VLYCSARTPNRHRAEALSIDHGECFOKPVL-----SHQICEPHGCGGVSVCPL 290
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 282 INVAKPEGQDNILERTL-----VYQSPNMGGSWTAHFTNRTIRMRNPGQALMK---IE 334
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 291 IPGG 294
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 335 EPGG 338

RESULT 14
ID Q59311 PRELIMINARY; PRT; 382 AA.
AC Q59311;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE EXO-ALPHA-SIALIDASE (EC 3.2.1.18).
GN NANH.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RA Chien C.H., Shann Y.J., Sheu S.Y.;
RT "Site-directed Mutations of the catalytic and conserved amino acids of
Neuraminidase gene, nanH.";
RL Enzyme Microb. Technol. 18:1-10(1996).
DR EMBL; X71123; CAA50436.1; -.
DR HSSP; P29768; 2SIL.
DR InterPro; IPR002860; BNR.
DR Pfam; PF02012; BNR; 4.
KW Hydrolase; Glycosidase.
SQ SEQUENCE 382 AA; 43028 MW; AFAF6471F7B2B431 CRC64;

Query Match
Best Local Similarity 5.1%; Score 120; DB 2; Length 382;
Matches 46; Conservative 25; Mismatches 63; Indels 36; Gaps 6;

QY 129 ICSQDAGYNSVDRLTEE--VIGPEVTHWATFVPGHG-IQLQSGRLIIPAYA----- 180
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 140 IYDDNGLTWSNKIDLTQSSKVKQPSNTIGWZGVGSGITWDDGTIVWPSQISAREN 199
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 ---YYIPFWFFCFRLPYRAPHSLMIYSDDLGLATWHHGRLLIKPMVTVCEVAEVIKAGH 237
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 200 ENNY-----SLIYSKDNGETWTMGNKVPNSNTSENVIEL-----D 237
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 238 PVLVCSARTP-NRHRAEALSIDHGECFO--KPVLSHQLCPEPHGCGGVS 284
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 238 VALIMSTRYDYSYRAAYISHDLGTTWEIYEPLNGKILTKGSGCGGSGFI 287
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
ID P77848 PRELIMINARY; PRT; 773 AA.
AC P77848;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SIALIDASE PRECURSOR (EC 3.2.1.18) (EXO-ALPHA-SIALIDASE)
DE (NEURAMINIDASE).
DE (NEURAMINIDASE).
GN NANH OR SIAH.
OS Clostridium tertium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1559;
RN [1]
RP SEQUENCE FROM N.A.
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 22.42 seconds  
(without alignments)  
466.287 Million cell updates/sec

Title: US-09-820-155-2  
Perfect score: 2348  
Sequence: 1 MEEVTSFSSPLFOQEDKR.....LSHVQDCSTPGMNSPKK 428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2348	100.0	428	US-09-423-340-2	Sequence 2, Appli
2	1916.5	81.6	428	US-09-423-340-4	Sequence 4, Appli
3	692	29.5	379	US-08-871-074-26	Sequence 26, Appli
4	201.5	8.6	626	5268290-2	Patent No. 5268290
5	110	4.7	51	US-08-871-074-11	Sequence 11, Appli
6	110	4.7	51	PCT-US94-05471-11	Sequence 11, Appli
7	98	4.2	931	US-08-936-135-16	Sequence 16, Appli
8	92	3.9	909	US-08-936-135-8	Sequence 8, Appli
9	91.5	3.9	782	US-09-146-283-4	Sequence 4, Appli
10	91.5	3.9	782	US-08-579-823A-4	Sequence 4, Appli
11	91.5	3.9	782	US-09-344-195-4	Sequence 4, Appli
12	91.5	3.9	926	US-08-936-135-14	Sequence 14, Appli
13	87.5	3.7	901	US-08-936-135-22	Sequence 22, Appli
14	87.5	3.7	914	US-08-936-135-12	Sequence 12, Appli
15	87	3.7	906	US-08-936-135-24	Sequence 24, Appli
16	86.5	3.7	1523	US-09-182-024A-2	Sequence 2, Appli
17	86.5	3.7	2471	US-08-185-432-16	Sequence 16, Appli
18	86.5	3.7	2471	US-08-083-590A-19	Sequence 19, Appli
19	86.5	3.7	2471	US-08-532-384-19	Sequence 19, Appli
20	85	3.6	794	US-07-885-972A-2	Sequence 2, Appli
21	85	3.6	794	US-07-885-972A-4	Sequence 4, Appli
22	85	3.6	794	US-08-865-203-2	Sequence 2, Appli
23	85	3.6	794	US-08-745-880-2	Sequence 2, Appli
24	85	3.6	794	US-08-745-880-4	Sequence 4, Appli
25	85	3.6	794	US-08-480-382-2	Sequence 2, Appli
26	85	3.6	794	US-08-480-382-4	Sequence 4, Appli
27	85	3.6	794	US-07-849-420-2	Sequence 2, Appli

28	85	3.6	794	4	US-09-253-854-2	Sequence 2, Appli
29	85	3.6	794	4	US-08-955-424-2	Sequence 2, Appli
30	85	3.6	909	3	US-08-936-135-18	Sequence 18, Appli
31	84.5	3.6	926	3	US-08-936-135-20	Sequence 20, Appli
32	84.5	3.6	1255	2	US-08-625-101-2	Sequence 2, Appli
33	84.5	3.6	1255	2	US-08-356-786-2	Sequence 9, Appli
34	84	3.6	484	2	US-08-252-493C-9	Sequence 9, Appli
35	84	3.6	484	3	US-09-276-197-9	Sequence 10, Appli
36	84	3.6	909	3	US-08-936-135-10	Sequence 6, Appli
37	83.5	3.6	1049	4	US-09-522-666-6	Sequence 8, Appli
38	83.5	3.6	1063	1	US-08-147-499A-8	Sequence 8, Appli
39	83.5	3.6	1063	1	US-08-482-847-8	Sequence 54, Appli
40	83.5	3.6	1141	1	US-08-131-365B-54	Sequence 54, Appli
41	83.5	3.6	1141	2	US-08-668-123-54	Sequence 4, Appli
42	83.5	3.6	2629	2	US-08-751-189-4	Sequence 4, Appli
43	83.5	3.6	2629	2	US-09-060-836-4	Sequence 4, Appli
44	83.5	3.6	2629	4	US-09-184-445-4	Sequence 4, Appli
45	82.5	3.5	1155	1	US-08-286-889-46	Sequence 46, Appli

ALIGNMENTS

RESULT 1  
US-09-423-340-2  
; Sequence 2, Application US/09423340  
; Patent No. 6225454  
; GENERAL INFORMATION:  
; APPLICANT: MITAGI, Taeko  
; APPLICANT: WADA, Tadashi  
; APPLICANT: YOSHIKAWA, Yuko  
; TITLE OF INVENTION: STALIDASE LOCALIZED IN PLASMA MEMBRANE AND  
; TITLE OF INVENTION: DNA CODING FOR THE SAME  
; FILE REFERENCE: 06699  
; CURRENT APPLICATION NUMBER: US/09/423,340  
; CURRENT FILING DATE: 1999-11-22  
; EARLIER APPLICATION NUMBER: JP 9-132174  
; EARLIER FILING DATE: 1997-05-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Bos primigenius  
US-09-423-340-2

Query Match	100.0%	Score	2348	DB	4	Length	428
Best Local Similarity	100.0%	Pred. No.	4.2e-251				
Matches	428	Conservative	0	Mismatches	0	Indels	0
Qy	1	MEEVTSFSSPLFOQEDKRGVTRIPALIIYVPPAHTFLAFAEKRSSKDEDLHLVLR	60				
Db	1	MEEVTSFSSPLFOQEDKRGVTRIPALIIYVPPAHTFLAFAEKRSSKDEDLHLVLR	60				
Qy	61	GLRTGQSVQWEPKSLMKATLPGRHTMPCPWERKSGYVYLFICVOGHTEROQIMSG	120				
Db	61	GLRTGQSVQWEPKSLMKATLPGRHTMPCPWERKSGYVYLFICVOGHTEROQIMSG	120				
Qy	121	RNPALRCFCISQDAGYSWSDVRLTEEVIGPEVTHWATFVGPCHGHIQLOSGRLIIPAYA	180				
Db	121	RNPALRCFCISQDAGYSWSDVRLTEEVIGPEVTHWATFVGPCHGHIQLOSGRLIIPAYA	180				
Qy	181	YYIPFWFCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEIGKAGHPVL	240				
Db	181	YYIPFWFCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEIGKAGHPVL	240				
Qy	241	YCSARTPNRRAEALSIDHGECFOKPVLSHOLCEPHCCQGSWSFCPLEIPGGCQDLAG	300				
Db	241	YCSARTPNRRAEALSIDHGECFOKPVLSHOLCEPHCCQGSWSFCPLEIPGGCQDLAG	300				
Qy	301	EDAPAIQCSPLLCSSVRPEPEAGTLESWLLYSHTPNKRRVLDLGIYLNQSPLEAACWSR	360				

Db 301 EDAPAIQSSPLLCSSVRPEPEAGTSLSEWLLYSHPTNKKRRVDLGIYLNQSPLEAACWSR 360  
QY 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFLFTDRILSHVQDCSTPG 420  
Db 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFLFTDRILSHVQDCSTPG 420  
QY 421 MNSEPSKK 428  
Db 421 MNSEPSKK 428

## RESULT 2

US-09-423-340-4  
; Sequence 4, Application US/09423340  
; Patent No. 6223454  
; GENERAL INFORMATION:  
; APPLICANT: MIYAGI, Taeko  
; APPLICANT: WADA, Tadashi  
; APPLICANT: YOSHIKAWA, YUKO  
; TITLE OF INVENTION: SIALIDASE LOCALIZED IN PLASMA MEMBRANE AND  
; TITLE OF INVENTION: DNA CODING FOR THE SAME  
; FILE REFERENCE: OP699  
; CURRENT APPLICATION NUMBER: US/09/423,340  
; CURRENT FILING DATE: 1999-11-22  
; EARLIER APPLICATION NUMBER: JP 9-132174  
; EARLIER FILING DATE: 1997-05-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-423-340-4

Query Match 81.6%; Score 1916.5; DB 4; Length 428;  
Best Local Similarity 82.4%; Pred. No. 2.2e-203;  
Matches 352; Conservative 32; Mismatches 40; Indels 3; Gaps 2;

QY 1 MEEVTSQSFSPFQDDKRGVYRIPALIVVPPAHTFLFAEKRSSKDEDALHLVLR 60  
Db 1 MEEVTSQSFSPFQDDKRGVYRIPALIVVPPAHTFLFAEKRSSKDEDALHLVLR 60  
QY 61 GLRTQSQVQWEPKLSMKATLPGHRTMPCVWERSGYVYLFVCGVHTERQQTMSG 120  
Db 61 GLRTQSQVQWEPKLSMKATLPGHRTMPCVWERSGYVYLFVCGVHTERQQTMSG 120  
QY 121 RNPARLCFCSQDAGYSDVRLTEEVIGPEVTHWATFVAGPGHGIQSGRLIIPAYA 180  
Db 121 RNPARLCFCSQDAGYSDVRLTEEVIGPEVTHWATFVAGPGHGIQSGRLIIPAYA 180  
QY 181 YYTPWFECFRLPYRARPHSLMIYSDDLGATWHHGRLLIKPMVTVCEVAVICKAGHPVL 240  
Db 181 YYTPWFECFRLPYRARPHSLMIYSDDLGATWHHGRLLIKPMVTVCEVAVICKAGHPVL 240  
QY 241 YCSARTPNRRHRAEALSDHGEQKPKVLSHQLCEPHGCGSVVSCFPLIPGCGODLAG 300  
Db 241 YCSARTPNRRHRAEALSDHGEQKPKVLSHQLCEPHGCGSVVSCFPLIPGCGODLAG 300  
QY 301 EDAPAIQSSPLLCSSVRPEPEAGTSLSEWLLYSHPTNKKRRVDLGIYLNQSPLEAACWSR 360  
Db 301 EDAPAIQSSPLLCSSVRPEPEAGTSLSEWLLYSHPTNKKRRVDLGIYLNQSPLEAACWSR 360  
QY 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFLFTDRILSHVQDCSTPG 420  
Db 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFLFTDRILSHVQDCSTPG 420  
QY 421 MNSEPSK 427  
Db 421 MNSEPSK 427

## RESULT 3

US-08-871-074-26  
; Sequence 26, Application US/08871074  
; Patent No. 5928915  
; GENERAL INFORMATION:  
; APPLICANT: Warner, Thomas C.  
; APPLICANT: Sliwowski, Mary B.  
; TITLE OF INVENTION: CHO Cell Sialidase By Recombinant DNA  
; TITLE OF INVENTION: Technology  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871,074  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/711,373  
; FILING DATE:  
; APPLICATION NUMBER: US/08/396,367  
; FILING DATE: 28-FEB-1995  
; APPLICATION NUMBER: 08/187327  
; FILING DATE: 25-JAN-1994  
; APPLICATION NUMBER: 08/062586  
; FILING DATE: 17-MAY-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Johnston, Sean A.  
; REGISTRATION NUMBER: 35,910  
; REFERENCE/DOCKET NUMBER: 826P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3562  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 379 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-871-074-26

Query Match 29.5%; Score 692; DB 2; Length 379;  
Best Local Similarity 38.6%; Pred. No. 5.1e-68;  
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;  
QY 13 LFQEDKRGVYRIPALIVVPPAHTFLFAEKRSSKDEDALHLVLRRLTQGS--VOW 70  
Db 13 LFQEDKRGVYRIPALIVVPPAHTFLFAEKRSSKDEDALHLVLRRLTQGS--VOW 70  
QY 71 EPLKSLMKATLPGHRTMPCVWERSGYVYLFVCGVHTERQQTMSGRLICFC 130  
Db 71 EPLKSLMKATLPGHRTMPCVWERSGYVYLFVCGVHTERQQTMSGRLICFC 130  
QY 69 QAEVTVQAYLEGRHSMSPCPDYDKQRTLFLFIAVRGOISHHQLOTVNVTLCIT 128  
Db 69 QAEVTVQAYLEGRHSMSPCPDYDKQRTLFLFIAVRGOISHHQLOTVNVTLCIT 128  
QY 131 SODAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQSGRLIIPAYVYIPWFF 188  
Db 131 SODAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQSGRLIIPAYVYIPWFF 188  
QY 129 STDHGKTSWAVQDLTDTTIGSTHODWATFVAGPGHGIQSGRLIIPAYVY 182  
Db 129 STDHGKTSWAVQDLTDTTIGSTHODWATFVAGPGHGIQSGRLIIPAYVY 182  
QY 189 CFRPL--YRAPHSLMIYSDDLGATWHHGRLLIKPMVTVCEVAVICKAGHPVLYCSART 246  
Db 189 CFRPL--YRAPHSLMIYSDDLGATWHHGRLLIKPMVTVCEVAVICKAGHPVLYCSART 246  
QY 183 --KQPIHAPAPSAFCFLSHDHGTSWELGHFVS--QNSLEQCAVEV--GTGAERVVYLNARS 238  
Db 183 --KQPIHAPAPSAFCFLSHDHGTSWELGHFVS--QNSLEQCAVEV--GTGAERVVYLNARS 238  
QY 247 PNRHRAEALSDHGEQKPKVLSHQLCEPHGCGSVVSCFPLIPGCGODLAGDAPAI 306  
Db 247 PNRHRAEALSDHGEQKPKVLSHQLCEPHGCGSVVSCFPLIPGCGODLAGDAPAI 306  
QY 239 CLGARVQAQSPNSGLDFQDNQVYVSKLVEPPKCGHGVIAF----- 278  
Db 239 CLGARVQAQSPNSGLDFQDNQVYVSKLVEPPKCGHGVIAF----- 278

```
QY 307 QQSPLLCSSVRPEACTLSESLLYSHPTNKKRRVLDGIYLNQSPLEAACWSRPWLHC 366
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 -----PNPTSADAL-DVLLYTHPTDSRKRNLGVLYLNQKPLDPTTWSAPILLAT 328
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 367 GPGCYSDLAAL-----ENEGFGLCEGCTKQCEQIAFRFLTDRILSHVOG 414
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 329 GICAYSDLQNMHGPGDGPQFGCLYE---SNVYEEIVFLMFLTKQAPPAVFG 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 4
5268290-2
; Patent No. 5268290
; APPLICANT: Hasegawa, Mamoru; Sakurada, Kazuhiro
; TITLE OF INVENTION: PROCESS FOR PRODUCING NEURAMINIDASE
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/384,709
; FILING DATE: 25-JUL-1989
; SEQ ID NO:2
; LENGTH: 626
5268290-2
Query Match 8.6%; Score 201.5; DB 6; Length 626;
Best Local Similarity 24.0%; Pred. No. 2e-13;
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;
QY 12 PLFQQED-----KRGV-TYRIPALIVYPHAHTFLAFKRSKSKDELDHLVL-RRLRT 64
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 12 PLYTEQDLAVNGREGFPNYRIPALT-VTPDGLLASVYDGRPTIDAPGPNILQRRSTDG 70
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 GOSVOWEPLAKMLKATLPGHRTNCPVWERKSGYVLFYICVQGHYTERQQIMSGRNP 124
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 GRTWGEQVVVSAGOTTAPIKGFSDPSLYVDRETGTIFNE-----HVYSQRQGFAGSRPG 124
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 125 -----RLCFTCSQDAGYSVSDVRDLTEEVIGPEVTHWATFVAGPGHGIOLQ----S 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 TDPADPNVLHANVATSDGGLTWSH-RTITAD-ITPD-PGWRSRFAASGEGIQLRVYCPHA 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 172 GRL-----ITPAVAYIPFWFFCRLPYRAPHSLMTIYSDDLGATWHHGLI----KPMV 222
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRLIQYTIINAAGAF-----QAVSVYSDDHGRTWRAGEAVGVGMDEK 225
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 TVCEVAEVIGKAGHPVLYCSARTPNRHRAEALSIDHGEFCQKPVLSHQLCEPHGCGS 282
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 TVELSDGRVLNLRD-----SAR--SGYRKVAVSTDGGHSGYGPVTIDRDLDPDTN--NAS 276
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 VVSFPCLEIPGGCQDLAGEDAPAIQOSPLLCSSVRPEPEAGTILSESLLYSHPTNKKRRV 342
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 277 IIRAFP-----DAP-----AGSARAKVLLFSNAASQTSR- 305
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 343 DLGIYLNQSPLEAAC-----WSRPWLHCPCGYSDLAALLENGLFGLCECCTKQCEQ 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 -----SQGTIRMSDGDGPWPKVQPGMSYSTLTALP-DGTGYLLYEPGT-----G 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 IAFRLFTDRILSHVOGDC---STPGCNSEPSKK 428
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 IRYANFN-----LAWLGGICAPFTIPDVALEPGQ 383
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 5
US-08-871-074-11
; Sequence 11, Application US/08871074
; Patent No. 5928915
; GENERAL INFORMATION:
; APPLICANT: Warner, Thomas C.
; APPLICANT: Sliwowski, Mary B.
; TITLE OF INVENTION: CHO Cell Stialidase By Recombinant DNA
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
```

```
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,074
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/711,373
FILING DATE:
APPLICATION NUMBER: US/08/396,367
FILING DATE: 28-FEB-1995
APPLICATION NUMBER: 08/187327
FILING DATE: 25-JAN-1994
APPLICATION NUMBER: 08/062586
FILING DATE: 17-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Johnston, Sean A.
REGISTRATION NUMBER: 35,910
REFERENCE/DOCKET NUMBER: 826P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3562
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-871-074-11
Query Match 4.7%; Score 110; DB 2; Length 51;
Best Local Similarity 47.9%; Pred. No. 5.5e-05;
Matches 23; Conservative 5; Mismatches 20; Indels 0; Gaps 0;
QY 116 QIMSGRNPRLCFCISQDAGYSVSDVRDLTEEVIGPEVTHWATFVAGCP 163
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 OLOTGVNVTLCRCHITSTDRKTSVAVODLTDTTIGSSDDQDXAXFVGVP 50
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 6
PCT-US94-05471-11
; Sequence 11, Application PC/TUS9405471
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: CHO Cell Stialidase By Recombinant DNA Technology
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05471
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/062586
```





Db 756 HGRILPSYDMEYQIVREGVIGK-----RSGEISID-----IRIS 792  
Qy 272 LCEPHGCGSVVSCFPLEIPGCGQDLAGEDAPAIQQSPLLCSSLVRPEAGTILSESLL 331  
Db 793 TDVPLENCEPIAFAD-EYEGDWSNSSSTSGA-----GDPSSGK-EKSWLY 838  
Qy 332 YSHP--TNKKRRVDGLVYLNQSPLEAACWSRPWTLHCGPCGYSLAA-----LEN 379  
Db 839 TLDPLITIAMSSGLVLLG-----ATC--AGLLLYC-TCSYGLSSRSCTTLEN 885

RESULT 9  
US-09-146-283-4  
: Sequence 4, Application US/09146283  
: Patent No. 5976546  
: GENERAL INFORMATION:  
: APPLICANT: Laus, Reiner  
: APPLICANT: Ruegg, Curtis L.  
: APPLICANT: Wu, Hongyu  
: TITLE OF INVENTION: Immunostimulatory Compositions  
: NUMBER OF SEQUENCES: 10  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Dehlinger & Associates  
: STREET: 350 Cambridge Ave. Suite 250  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94306  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/146,283  
: FILING DATE: 03-SEPT-1998  
: CLASSIFICATION: 536  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Judge, Linda R.  
: REGISTRATION NUMBER: 42,702  
: REFERENCE/DOCKET NUMBER: 7636-0010.21  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 650-324-0880  
: TELEFAX: 650-324-0960  
: INFORMATION FOR SEQ ID NO: 4:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 782 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: HYPOTHETICAL: NO  
: ORIGINAL SOURCE:  
: ORGANISM: homo sapiens  
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
US-09-146-283-4

Query Match 3.9%; Score 91.5; DB 2; Length 782;  
Best Local Similarity 20.7%; Pred. No. 0.41;  
Matches 62; Conservative 27; Mismatches 112; Indels 99; Gaps 13;  
Qy 94 ERKSGYVLF--ICVQGHVTERQQIMSGR---NPARLCFICSDAGYSWSDVRLTTEE 147  
Db 404 EETGYLYISAWPDSLPDLVSFQNLQVIRGILHNGA--YSLTLQGLGISWLGRLSREL 461  
Qy 148 VIGPEVTHWATFVAGPGHGIGLOSGRLIIPAYAYIIPWFECFRLPYRAPHSLMIYSDD 207  
Db 462 GSGGLALIH-----HNTHL-----CFVHTVP-WDOLFNRPHOALLHTANRPEDE 503  
Qy 208 L-----GATWHHGRLLKPMVTVCE-----VAEVIKAGHPVLYCSAR- 245  
Db 504 CVGEGLAGHQLCARGHCWGP-----PTQCVNCSQFLRQECVEECRVLQGLPREYVNAH 559

Qy 246 -----TPNRHRAEALSIDHGECFQ-----KPVLSHQ----- 271  
Db 560 CLPCHPCQFQNGSVTCFGEADQCVACAHYKDPFPCVARGPCSVKPDLSLMPWKFPDE 619  
Qy 272 --LCEPHGCGSVVSCFPLEIPGCGQDLAGEDAPAIQQSPLLCSSLVRPEAGTILSES 329  
Db 620 EGACQP-----CPINCTHSCVDLDDKGCPCQASPLTSLAPARSPSPSTQPW 668  
RESULT 10  
US-08-579-823A-4  
: Sequence 4, Application US/08579823A  
: Patent No. 6080409  
: GENERAL INFORMATION:  
: APPLICANT: Laus, Reiner  
: APPLICANT: Ruegg, Curtis L.  
: APPLICANT: Wu, Hongyu  
: TITLE OF INVENTION: Immunostimulatory Composition and Method  
: NUMBER OF SEQUENCES: 10  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Dehlinger & Associates  
: STREET: 350 Cambridge Ave. Suite 250  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94306  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/579,823A  
: FILING DATE: 03-DEC-1998  
: CLASSIFICATION: 536  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Judge, Linda R.  
: REGISTRATION NUMBER: 42,702  
: REFERENCE/DOCKET NUMBER: 7636-0010  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 650-324-0880  
: TELEFAX: 650-324-0960  
: INFORMATION FOR SEQ ID NO: 4:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 782 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: HYPOTHETICAL: NO  
: ORIGINAL SOURCE:  
: ORGANISM: homo sapiens  
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
US-08-579-823A-4

Query Match 3.9%; Score 91.5; DB 3; Length 782;  
Best Local Similarity 20.7%; Pred. No. 0.41;  
Matches 62; Conservative 27; Mismatches 112; Indels 99; Gaps 13;  
Qy 94 ERKSGYVLF--ICVQGHVTERQQIMSGR---NPARLCFICSDAGYSWSDVRLTTEE 147  
Db 404 EETGYLYISAWPDSLPDLVSFQNLQVIRGILHNGA--YSLTLQGLGISWLGRLSREL 461  
Qy 148 VIGPEVTHWATFVAGPGHGIGLOSGRLIIPAYAYIIPWFECFRLPYRAPHSLMIYSDD 207  
Db 462 GSGGLALIH-----HNTHL-----CFVHTVP-WDOLFNRPHOALLHTANRPEDE 503  
Qy 208 L-----GATWHHGRLLKPMVTVCE-----VAEVIKAGHPVLYCSAR- 245  
Db 504 CVGEGLAGHQLCARGHCWGP-----PTQCVNCSQFLRQECVEECRVLQGLPREYVNAH 559  
Qy 246 -----TPNRHRAEALSIDHGECFQ-----KPVLSHQ----- 271

Db 560 CLPCHPECQONGSVTCFGPEADQCVACAHYKDPFCVARCPGKVPDLISYMPDIWKFPDE 619  
QY 272 --LCEPHGCGQSVWFCPLIEPGGQDLAGEDAPAIQQSPILCSSLVRPEAGTLESW 329  
Db 620 EGACQP-----CPINCTHSCVDLDDKCPAEQASPLTSLEAPARSPSPSTQPW 668

RESULT 11  
US-09-344-195-4  
; Sequence 4, Application US/09344195  
; Patent No. 6210662  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; Ruegg, Curtis L.  
; Wu, Hongyu  
; TITLE OF INVENTION: Immunostimulatory Compositions  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/344,195  
; FILING DATE: 24-Jun-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/146,283  
; FILING DATE: 03-SEPT-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Judge, Linda R.  
; REGISTRATION NUMBER: 42,702  
; REFERENCE/DOCKET NUMBER: 7636-0010.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-0880  
; TELEFAX: 650-324-0960  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 782 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: homo sapiens  
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-344-195-4

Query Match 3.9%; Score 91.5; DB 4; Length 782;  
Best Local Similarity 20.7%; Pred. No. 0.41;  
Matches 62; Conservative 27; Mismatches 112; Indels 99; Gaps 13;

QY 94 ERKSGVYLFF---ICVQGHVTERQIMSGR---NPARLCFCISQDAGYSWSDVRDLTEE 147  
Db 404 EEITGYLYISAWPDLPLDSVFQNLQVIRGRILHNGA--YSLTQGLGSLGLRLSLREL 461  
QY 148 VIGPEYTHWATFAGVGHGIGLOSGRLIIPAYAYIIPFWFFCFRLPYRAPHSLMTYSDD 207  
Db 462 GSGLALIH-----HNTHL-----CFVHTVP-WDQLFRNPHQALLHTANRPEDE 503  
QY 208 L-----GATWHGRLIKRPMVTECE-----VAEVIGKAGHPVLYCSAR- 245  
Db 504 CVGEGLACHQLCARGHCWGP-----PTQCVCNSQFLRGQCEVCECRVLQGLPREYVNAH 559

QY 246 -----TPNRHRAEALSIDHGECFQ-----KPVLSHQ----- 271  
Db 560 CLPCHPECQONGSVTCFGPEADQCVACAHYKDPFCVARCPGKVPDLISYMPDIWKFPDE 619  
QY 272 --LCEPHGCGQSVWFCPLIEPGGQDLAGEDAPAIQQSPILCSSLVRPEAGTLESW 329  
Db 620 EGACQP-----CPINCTHSCVDLDDKCPAEQASPLTSLEAPARSPSPSTQPW 668

RESULT 12  
US-08-936-135-14  
; Sequence 14, Application US/08936135  
; Patent No. 6054293  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: He, Zhigang  
; TITLE OF INVENTION: Semaphorin Receptors  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 75 DENISE DRIVE  
; CITY: HILLSBOROUGH  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94010  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/936,135  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC97-288-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 343-4341  
; TELEFAX: (650) 343-4342  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 926 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-936-135-14

Query Match 3.9%; Score 91.5; DB 3; Length 926;  
Best Local Similarity 22.9%; Pred. No. 0.54;  
Matches 55; Conservative 30; Mismatches 90; Indels 65; Gaps 13;

QY 170 QSGRLIIPAYAYIIPFWFFCFRLPYRA-----RPHSLMTYSDDLGATWH 213  
Db 698 QYGRLLSP--PVHLPRSPVCMFEFYQAMGGHVALQVREASQESKLLWVIREDDQSEWK 755  
QY 214 HGRLLKPMVTVCEVA--EVIKAGHPVLYCSARTNRRHRAEALSIDHGECFQKPVLSHQ 271  
Db 756 HGRILLPSYDMEYQIVFEGVICKG-----RSGEISGDD-----IRIS 792  
QY 272 LCEPHGCGQSVWFCPLIEP-----GGQDLAGEDAPAIQQSPILCSSLVRPEAGTLES 326  
Db 793 TDVPLENCMEPIISAFV-DIPETHGGEYDEIDDEYEGDWSNSSSTSGAGDPSSGK-E 850  
QY 327 ESKLLYSHP--INKRRVDLGIYLNOSPLEAACWSRPWILHCGPCGYSDLAA-----LEN 379  
Db 851 KSWLYTLDPILITIIMSSGLVLLG-----ATC--AGLLLYC-TCSYSLGSSRSCTTLEN 902

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	3.7%	Score 87.5;	DB 3;	Length 901;
	Best Local Similarity	22.5%	Pred. No. 1.4;		
	Matches	38;	Conservative	19;	Mismatches 57; Indels 55; Gaps 7;
Qy	170	QSGRLIIPAVAYVYTPWFEECFERLPYRA-----RPHSLMIYSDDLGATWH	213		
			:		
		:	:		
		:	:		
		:	:		
Db	698	OYGLRLISP--PVHLPRSPVCMCFQYQAMGGHGVALQVVVREASQESKLLWVIREDOGSEWK	755		
Qy	214	HGRLLIKPMVTVCEBSVA--EVIGKAGHPVLYCSARTPNRHRAEALSDHGECFKQKPVLSHQ	271		
			:		
		:	:		
		:	:		
Db	756	HGRLLPSYDMEYQIVFEGVICKG-----RSGETSIDD-----IRIS	792		
Qy	272	LCEPHHCQGSWSFCFLEIPGGCODLAGEDAPAIQQSPLLCSSVRRPEP	320		
		:	:		
		:	:		
		:	:		
Db	793	TDVPLENCMPISAFAGTLPPTGE-----PWTVDVTP-----VQPTP	829		

```

RESULT 14
US-08-936-135-12
: Sequence 12, Application US/08936135
: Patent No. 6054293
: GENERAL INFORMATION:
: APPLICANT: Tessier-Lavigne, Marc
: APPLICANT: He, Zhigang
: APPLICANT: Chen, Hang
: TITLE OF INVENTION: Semaphorin Receptors
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

```

```

; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 914 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-936-135-12

Query Match          3.7%; Score 87.5; DB 3; Length 914;
Best Local Similarity 22.8%; Pred. No. 1.5;
Matches 55; Conservative 26; Mismatches 81; Indels 79; Gaps 13;

Qy 170 QSGRLIPAYAYITPFFFCRRLPYRA-----RPHSLMIYSDDLGATWH 213
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 698 QYGRLLIS--PVHLPRSPVCMEFOYQAMGGHGVALQVYVREASQESKLLWLVIREDOGSWK 755
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Qy 214 HGLRIKPMTVECEVA--EVIGK-----AGHPVLVCSARTPNRHRAEALSIDHGECFOK 265
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 756 HGRILLPSYDWEYQIVFEGVIKGSRGEISODDI--RTSTDVPLENCMEPISAFAGEDFKD 814
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Qy 266 PVLSHQLCEPPHGCQGVSVFPCLEIPGCCODLAGEDAPAIQQSPLLCSSVRPEAGTTL 325
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 815 -----EYEGDWSNSSSTSGA-----GDPSSGK- 837
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Qy 326 SESWLVSHP--TNKKRRVDLGYLNOSPLBAACSWRPWIHCGPGCYSDLA-----LE 378
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 838 EKSWLYTLDPLITTIAMSSLGVLIG-----ATC--AGLLLYC-TCSYSLSSRSCTTLE 889
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Qy 379 N 379
Db 890 N 890

RESULT 15
US-08-936-135-24
; Sequence 24, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-936-135-24

```

```

Query Match      3.7%; Score 87; DB 3; Length 906;
Best Local Similarity 23.1%; Pred. No. 1.6;
Matches 39; Conservative 18; Mismatches 62; Indels 50; Gaps 7;

QY 170 QSGRLIIPAYYIPFWFFCFRLPYRA-----RPHSLMIYSDDLGATWH 213
      ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 698 QYGRLLISP--PVHLPRSPVCMFQYQAMGGHGVALQVVRASQESKLLWVIREQGSWK 755

QY 214 HGRLLKPMWTVCEVA--EVIGKAGHPVLYCSARTPNRHRRAELSDHGECEKPKVLSHQ 271
      ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 756 HGRILLPSYDMEYQIVFEGVIGKG-----RSGEISIDD-----IRIS 792

QY 272 LCEPPHGCQGSVVSFCPLPTPGCQDLAGEDAPAIQQSPLLCSSVRPEP 320
      | : : : : : : : : : : : : : : : : : : : : :
Db 793 TDVPLENCMEPISAFAGEDFKGG--TLPPGTETPTVTVP-----VQPIP 834

```

Search completed: October 7, 2002, 14:59:11  
Job time: 33 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 211.59 Seconds  
(without alignments)  
711.979 Million cell updates/sec

Title: US-09-820-155-2  
Perfect score: 2348  
Sequence: 1 MEEVTSCTSSFLFQEDKR.....LSHVQGDCTPGMNSERSKK 428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 3502263 seqs, 351980561 residues  
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1:	/cgn2_6/ptodata/2/paa/PTCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US100_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US101_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2348	100.0	428	22	US-09-820-155-2
2	1916.5	81.6	428	22	US-09-820-155-4
3	875	37.3	161	22	US-09-820-155-25
4	653	27.8	372	19	US-09-556-870A-2
5	520	22.1	362	26	US-60-207-359-294
6	476.5	20.3	240	22	US-09-820-155-26
7	295	12.6	245	1	PCT-US01-01240-372

8	295	12.6	245	1	PCT-US01-01324-1918	Sequence 1918, Ap
9	295	12.6	245	21	US-09-764-872-372	Sequence 372, App
10	295	12.6	245	21	US-10-079-379-1918	Sequence 1918, Ap
11	255.5	10.9	425	21	US-09-758-442-464	Sequence 464, App
12	213	9.1	151	1	PCT-US01-01354-18118	Sequence 18118, A
13	213	9.1	151	21	US-09-764-905-18118	Sequence 18118, A
14	213	9.1	151	24	US-10-092-399-18118	Sequence 18118, A
15	177	7.5	413	18	US-09-417-507-38401	Sequence 38401, A
16	147.5	6.3	937	19	US-09-583-110-3166	Sequence 3166, Ap
17	147.5	6.3	975	15	US-09-107-433-3032	Sequence 3032, Ap
18	140.5	6.0	683	24	US-10-091-007-64	Sequence 64, Appl
19	138	5.9	248	1	PCT-US00-05883-958	Sequence 958, App
20	138	5.9	248	23	US-09-935-299-958	Sequence 958, App
21	137.5	5.9	394	21	US-09-738-626-5219	Sequence 5219, Ap
22	110	4.7	51	12	US-08-871-076-9	Sequence 9, Appli
23	108.5	4.6	642	13	US-08-911-393-4	Sequence 4, Appli
24	108.5	4.6	642	23	US-09-955-909-4	Sequence 4, Appli
25	106.5	4.5	669	24	US-10-086-913-2	Sequence 2, Appli
26	106.5	4.5	1060	13	US-08-911-393-2	Sequence 2, Appli
27	106.5	4.5	1060	23	US-09-955-909-2	Sequence 2, Appli
28	105.5	4.5	666	21	US-09-745-008-2	Sequence 2, Appli
29	104.5	4.5	763	1	PCT-US01-08631-3127	Sequence 3127, A
30	102	4.3	567	18	US-09-488-725A-1803	Sequence 1803, Ap
31	101	4.3	571	21	US-09-760-466-907	Sequence 907, App
32	98	4.2	931	1	PCT-US98-14290-16	Sequence 16, Appl
33	98	4.2	931	18	US-09-439-711-16	Sequence 16, Appl
34	97	4.1	583	21	US-09-760-466-1379	Sequence 1379, Ap
35	97	4.1	925	15	US-09-116-473-2	Sequence 2, Appli
36	96	4.1	347	26	US-60-242-679-1068	Sequence 1068, Ap
37	96	4.1	391	24	US-10-015-127-10231	Sequence 10231, A
38	94.5	4.0	596	18	US-09-488-725A-5375	Sequence 5375, Ap
39	94.5	4.0	2820	20	US-09-614-150-16680	Sequence 16680, A
40	94	4.0	820	26	US-60-150-584-690	Sequence 690, App
41	93	4.0	1078	26	US-60-207-315-428	Sequence 428, App
42	93	4.0	1078	26	US-60-212-659-523	Sequence 523, App
43	93	4.0	1078	26	US-60-230-435-1751	Sequence 1751, Ap
44	92	3.9	476	24	US-10-029-386-32293	Sequence 32293, A
45	92	3.9	479	1	PCT-US01-07143-44	Sequence 44, Appl

ALIGNMENTS

RESULT 1  
US-09-820-155-2  
; Sequence 2, Application US/09820155  
; GENERAL INFORMATION:  
; APPLICANT: Miyagi, Taeko  
; APPLICANT: Wada, Tadashi  
; APPLICANT: Yoshikawa, Yuko  
; TITLE OF INVENTION: Stalidase Localized in the Plasma Membrane  
; FILE REFERENCE: 20111-0046 (45455-252923)  
; CURRENT APPLICATION NUMBER: US/09/820,155  
; CURRENT FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 09/423,340  
; PRIOR FILING DATE: 1999-11-22  
; PRIOR APPLICATION NUMBER: PCT/JP98/02072  
; PRIOR FILING DATE: 1998-05-11  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Bos primigenius taurus  
US-09-820-155-2

Query Match 100.0%; Score 2348; DB 22; Length 428;  
Best Local Similarity 100.0%; Pred. No. 1.6e-217;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEEVTSCTSSFLFQEDKRGVYRIPALTYVPATLAFAEKRSSKDEDAHLVLRR 60  
|||||

Db 1 MEEVTSFSPFOEDKRGVYTRIPALIIYVPPAHTLAFAEKRSSKDEALHLVLR 60  
QY 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIIMG 120  
Db 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIIMG 120  
QY 121 RNPARLCFICSODAGYSWSDVRDLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180  
Db 121 RNPARLCFICSODAGYSWSDVRDLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180  
QY 181 YYIPWFCCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEVTGRAGHPVL 240  
Db 181 YYIPWFCCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEVTGRAGHPVL 240  
QY 241 YCSARTPNRRAEALSIDHGECFQKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300  
Db 241 YCSARTPNRRAEALSIDHGECFQKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300  
QY 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHTPNKKRRVLDLGIYLNQSPLEAACWSR 360  
Db 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHTPNKKRRVLDLGIYLNQSPLEAACWSR 360  
QY 361 PWILHCGPCGYSDLAALNEGLFGCLFECGTQKQCEQIAFRLFTDREILSHVQDCSTPG 420  
Db 361 PWILHCGPCGYSDLAALNEGLFGCLFECGTQKQCEQIAFRLFTDREILSHVQDCSTPG 420  
QY 421 MNSEPSK 428  
Db 421 MNSEPSK 428

## RESULT 2

US-09-820-155-4  
; Sequence 4, Application US/09820155  
; GENERAL INFORMATION:  
; APPLICANT: Miyagi, Taeko  
; APPLICANT: Wada, Tadashi  
; APPLICANT: Yoshikawa, Yuko  
; TITLE OF INVENTION: Sialidase Localized in the Plasma Membrane  
; FILE REFERENCE: 20111-0046 (45455-252923)  
; CURRENT APPLICATION NUMBER: US/09/820,155  
; CURRENT FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 09/423,340  
; PRIOR FILING DATE: 1999-11-22  
; PRIOR APPLICATION NUMBER: PCT/JP98/02072  
; PRIOR FILING DATE: 1998-05-11  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-820-155-4

Query Match 81.6%; Score 1916.5; DB 22; Length 428;  
Best Local Similarity 82.4%; Pred. No. 8e-176;  
Matches 352; Conservative 32; Mismatches 40; Indels 3; Gaps 2;

QY 1 MEEVTSFSPFOEDKRGVYTRIPALIIYVPPAHTLAFAEKRSSKDEALHLVLR 60  
Db 1 MEEVTSFSPFOEDKRGVYTRIPALIIYVPPAHTLAFAEKRSSKDEALHLVLR 60  
QY 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIIMG 120  
Db 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIIMG 120  
QY 121 RNPARLCFICSODAGYSWSDVRDLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180  
Db 121 RNPARLCFICSODAGYSWSDVRDLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180  
QY 181 YYIPWFCCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEVTGRAGHPVL 240  
Db 181 YYIPWFCCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEVTGRAGHPVL 240

Db 181 YYIPWFCCFQPCCKTRPHSLMIYSDDLGVTWHHGRLLRPMVTVCEVAEVTGRAGHPVL 240  
QY 241 YCSARTPNRRAEALSIDHGECFQKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300  
Db 241 YCSARTPNRRAEALSIDHGECFQKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300  
QY 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHTPNKKRRVLDLGIYLNQSPLEAACWSR 360  
Db 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHTPNKKRRVLDLGIYLNQSPLEAACWSR 360  
QY 361 PWILHCGPCGYSDLAALNEGLFGCLFECGTQKQCEQIAFRLFTDREILSHVQDCSTPG 420  
Db 361 PWILHCGPCGYSDLAALNEGLFGCLFECGTQKQCEQIAFRLFTDREILSHVQDCSTPG 420  
QY 421 MNSEPSK 427  
Db 420 RN--PSQ 424

## RESULT 3

US-09-820-155-25  
; Sequence 25, Application US/09820155  
; GENERAL INFORMATION:  
; APPLICANT: Miyagi, Taeko  
; APPLICANT: Wada, Tadashi  
; APPLICANT: Yoshikawa, Yuko  
; TITLE OF INVENTION: Sialidase Localized in the Plasma Membrane  
; FILE REFERENCE: 20111-0046 (45455-252923)  
; CURRENT APPLICATION NUMBER: US/09/820,155  
; CURRENT FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 09/423,340  
; PRIOR FILING DATE: 1999-11-22  
; PRIOR APPLICATION NUMBER: PCT/JP98/02072  
; PRIOR FILING DATE: 1998-05-11  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 25  
; LENGTH: 161  
; TYPE: PRT  
; ORGANISM: Bos sp.  
US-09-820-155-25

Query Match 37.3%; Score 875; DB 22; Length 161;  
Best Local Similarity 98.8%; Pred. No. 1e-75;  
Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 49 KDEDALHLVLRRLTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQ 108  
Db 1 KDEDALHLVLRRLTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQ 60  
QY 109 GHVTERQIIMSGRNPARLCFICSODAGYSWSDVRDLTEEVIGPEVTHWATFVAVGPGHGIQ 168  
Db 61 GHVTERQIIMSGRNPARLCFICSODAGYSWSDVRDLTEEVIGPEVTHWATFVAVGPGHGIQ 120  
QY 169 LOSGRLLIIPAYAYIPWFCCFRLPYRAPHSLMIYSDDLG 209  
Db 121 LOSGRLLIIPAYAYIPWFCCFRLPYRAPHSLMIYSDDLG 161

## RESULT 4

US-09-556-870A-2  
; Sequence 2, Application US/09556870A  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Robert  
; APPLICANT: Chris, Fronda  
; APPLICANT: Zeng, Guichao  
; TITLE OF INVENTION: CONA SEQUENCE OF MOUSE BRAIN SIALIDASE GENE  
; FILE REFERENCE: SEQUENCE (02940110AA)  
; CURRENT APPLICATION NUMBER: US/09/556,870A  
; CURRENT FILING DATE: 2000-04-21  
; PRIOR APPLICATION NUMBER: US 60/130,712  
; PRIOR FILING DATE: 1999-04-22











PRIOR APPLICATION NUMBER: 60/249,218  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,208  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,213  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,212  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,207  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,245  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,244  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,217  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,211  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,215  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,264  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,214  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,297  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/232,400  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/231,242  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/232,081  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/232,080  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/231,414  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/231,244  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/233,064  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/233,063  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232,397  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232,399  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232,401  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/241,808  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241,826  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241,786  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241,221  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/246,475  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/231,243  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/233,065  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232,398

Query Match  
Best Local Similarity 9.1%; Score 213; DB 21; Length 151;  
Matches 40; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MEEVTCSPFLFOQEDRGVYTRIPALYYVPPAHTFLAFAEKRSSSKDEDAHLV 57  
DB 76 MEEVTCSPFLFOQEDRGVYTRIPALYYXPTXXFLXFAEKSTRDEDAHLV 132

RESULT 14  
US-10-092-399-18118  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, proteins, and Antibodies  
; FILE REFERENCE: PC004C1  
; CURRENT APPLICATION NUMBER: US/10/092,399  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 42506  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18118  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (99)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: misc\_feature  
; LOCATION: (108)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: misc\_feature  
; LOCATION: (111)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: misc\_feature  
; LOCATION: (112)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: misc\_feature  
; LOCATION: (115)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: misc\_feature  
; LOCATION: (130)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-092-399-18118

Query Match  
Best Local Similarity 9.1%; Score 213; DB 24; Length 151;  
Matches 40; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MEEVTCSPFLFOQEDRGVYTRIPALYYVPPAHTFLAFAEKRSSSKDEDAHLV 57  
DB 76 MEEVTCSPFLFOQEDRGVYTRIPALYYXPTXXFLXFAEKSTRDEDAHLV 132

RESULT 15  
US-09-417-507-38401  
; GENERAL INFORMATION:  
; APPLICANT: KEITH G. WEINSTOCK ET AL.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS  
; FILE REFERENCE: PATH99-10  
; CURRENT APPLICATION NUMBER: US/09/417,507  
; CURRENT FILING DATE: 1999-10-14  
; NUMBER OF SEQ ID NOS: 44312  
; SEQ ID NO 38401  
; LENGTH: 413  
; TYPE: PRT  
; ORGANISM: A.fumigatus  
US-09-417-507-38401

Query Match  
Best Local Similarity 7.5%; Score 177; DB 18; Length 413;  
Matches 63; Conservative 33; Mismatches 61; Indels 40; Gaps 13;  
QY 12 PLFQOED-----KRGV---TYRIPALYYVPPAHTFLAFAE-KRSSSKDEDAHLVLR 60  
DB 43 PLFRSANMSPDKLSTGTGTFHGFHFRIPAVVVTGTR-ILAFAEGRRTNQDFGDLNLYK 101

QY 61 GLRTGQS---VQWEPKSLMKATLPGHRTM-NPCPVWERKSGYVYLF----- 104  
Db 102 TKTTANNGASPSDWEPLREVGS---GAGTWCNPTPVVD-DDNTIYLFSLWNGATYSQNG 157  
QY 105 --ICVQGHVTER-QOIMSGRNPRLCFTICSDODAGYSWSDVRDLTEEVIGPEVTHWATFAV 161  
Db 158 KDVLPDGTVTKKIDSTWEGRR--HLYLTESRDDGNTWSKPVDLTKE-LTPD--GWAWDAV 212  
QY 162 GPGHGIQLOSGRLIIPA 178  
Db 213 GPGNGIRLTGTGELVIPA 229

Search completed: October 7, 2002, 15:04:31  
Job time: 353 sec

34 MEEVTCSENSPLFRQEDDRGT



```
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27331
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-27331

Query Match          30.3%; Score 711.5; DB 5; Length 380;
Best Local Similarity 40.7%; Pred. No. 4.4e-55;
Matches 170; Conservative 59; Mismatches 114; Indels 75; Gaps 16;

QY 10 SSPLFOED--KRGV--TYRIPALIVVPPAHTFLAFAEKRRSSKDEDALHLVLRG--LR 63
DB 3 SLPLVKESVFQSGAHAYRIPALILYLPQQSLLAFAEQRAKKDEHAELIVLRGDDYDAP 62
QY 64 TGOVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFICVQGHVTERQOIMSGRNP 123
DB 63 THQ-VQMQAEVVAQARLDGHRSMNCPPLYDAQGTGLFLEFFIAIPGVTEQOQLOTRANV 121
QY 124 ARLCFTSQDAGYSWSDVRLTEEVIGPEVTHWATFVAGFGHGIGQL--QSGRLIIPAYAY 181
DB 122 TRLCQVSTDHGRTWSRDLTDAAGIPAYRWSTFVAGFGHCLQNDRAARSLVWPAYAY 181
QY 182 Y-----IPWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEIVG 233
DB 182 RKLHPTQRPIS-AFCF-----LSHDGRTWARGHFV-AQDTLECOVAEV-- 224
QY 234 KAGHP-VLYCSARTPNRHRAEALSIDHGECFQKPVLSHOLCE--PPHGCQSGSVVFCPL 291
DB 225 ETGEQVRVTLNARSHLRARVQAGSTNDGLDFQESQLVKLVPPPOGCGSVISF----- 279
QY 292 PGCQDLAGEDAPAIQOOSPLLCSSVRPEPEAGTLSSES-WLLYSHPTNKKRRVDLGIYLNQ 350
DB 280 -----PSPRSGPGSPAQWLLYTHPTSHWQRADLGAYLNP 313
QY 351 SPLEAACWSPWILHCGPCGYSDLAAL----ENEGLFGCLFECGTQKECEQIAFRFLT 404
DB 314 RPPAPEWSEPVLLAKGSCAYSDLQSMGTGPDGSPFLFGCLYEIA---NDYEEIVFLMFT 368

RESULT 5
US-09-791-537-15522
; Sequence 15522, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15522
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-15522

Query Match          30.0%; Score 705; DB 5; Length 379;
Best Local Similarity 39.8%; Pred. No. 1.7e-54;
Matches 164; Conservative 63; Mismatches 117; Indels 68; Gaps 14;

; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27331
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-27331

Query Match          29.5%; Score 692; DB 5; Length 379;
Best Local Similarity 38.6%; Pred. No. 2.4e-53;
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;

QY 13 LFOQDKRGVTRIPALIVVPPAHTFLAFAEKRRSSKDEDALHLVLRRLGTQGS--VOW 70
DB 12 LFOQGD---YAYRIPALIVLSKQKTLAFAEKRLTKTDEHADLFVLRGSSYNADTHQVQW 68
QY 71 EPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFICVQGHVTERQOIMSGRNPRLCFTIC 130
DB 69 QAEVTVQAYLEGRHSMSPCPLYDKQTRTLFLFFIAVGOISHHQLQTVGNVTRLCBIT 128
QY 131 SQDAGYSWSDVRLTEEVIGPEVTHWATFVAGFGHGIGIQLO--SGRLIIPAYAYIIFWFF 188
DB 129 STDHGKTWSAVQDLTDTTIGSTHQDWATFVGFGHCLQRLNRTAGSLVPAAYR----- 182
QY 189 CFRPLP--YRARPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEIVGKAGHPVLYCSART 246
DB 183 --KQPIHAPAPSAFCFLSHDHSSTWELGHFVS-QNSLECOVAEV-GTGAERVVYLNARS 238
QY 247 PNHRRAEALSIDHGECFQKPVLSHOLCEPPHGCQSGSVVFCPLIFPGCQDLAGEDAPAI 306
```

Db 239 CIGARVQAQSPNSGLDFQDNQVSKLVPPKCHGSGVIAF----- 278  
Qy 307 QOSPLLCSSVRPEAGTSLSEWLLYSHTNKKRRVDLGIYLNOSPLEAACWSRPWILHC 366  
Db 279 -----PNPTSKADAL-DWLLYIHTDSRKTNLGVYLNQKPLDPTTWSAPTLLAT 328  
Qy 367 GPCGYSDLAAL-----ENGLGFCCLFECGCKQCEQIAFRFLTDRILSHVOG 414  
Db 329 GICAYSDLQNHGDPGSPQSGCLYE---SNNYEIVELMTLTKQAFPAVFG 377

RESULT 7  
US-09-791-537-87692  
; Sequence 87692, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 87692  
; LENGTH: 415  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-87692

Query Match 10.9%; Score 255.5; DB 5; Length 415;  
Best Local Similarity 28.0%; Pred. No. 2.9e-14;  
Matches 116; Conservative 46; Mismatches 130; Indels 123; Gaps 23;

Qy 23 TYRIPALIVPPAHTFLAFAEKRS--SSKDEDALHLVLRRLTQSGVOWEPLKSLMK-AT 80  
Db 76 TFRIP-LITATPRGTLTLLAFAEARKMSSDEGAKFTALRRSMDQGST--WSPATFIVNDGD 132  
Qy 81 LPHGHTMPCPVWERKSGVYVLF--ICVQGHVTERQQIMSGRNPARLCFICSDAGYSWS 139  
Db 133 VPDGLNLG-AVVSDETGVVFLFYSLCAH-----KAGCOVASTMLVWSKDDGVWS 182  
Qy 140 DVRLTEVIGPEVTHWATFVAGVGHGIGLO-----SGLIIPAYAYIIPFFWFCFLPYR 195  
Db 183 TPRNLSLD-IGTEV-----FAPGPGSGIQKQREPRKGLIIVCGHGTLEDRDGVFC----- 230  
Qy 196 ARPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIKAGHPVLYCSARTPNRRAEAL 255  
Db 231 -----LLSDDHGASWRYGS-----GVSGIP----- 250  
Qy 256 SIDHGEQFKVLSHQLECPHPGCGSVV-----SFCPLEIPGGCQDLAGEDAPAIQ 307  
Db 251 ---YGQPKQENDFPDECPYELPDGVSVINARNQNNYHCHCRIV-----LRSYDA----- 298  
Qy 308 QSPILCSSVRP-----BPE-----AGTL--SESWLLYSHTNKKRRVDLGIYLNOSPL 353  
Db 299 -----CDTLRPRDVTDFPELDPVVAAGAVVTSSGIVFFSNPAHPFVRNLI--TLRWSFS 351  
Qy 354 EAACWSRPWI-LHCGPCGYSDLAALENEGLFG-----CLFECGCKQCEQIA 399  
Db 352 NGTSWRKETVOLWPGSPGSSYSLATLEG-SMDGEGEQAPOLYVLYEKGRRNHHTESIS 405

RESULT 8  
US-10-217-623-464  
; Sequence 464, Application US/10217623  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PM031C1N  
; CURRENT APPLICATION NUMBER: US/10/217,623

; CURRENT FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 09/758,442  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 710  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 464  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-217-623-464

Query Match 10.9%; Score 255.5; DB 6; Length 425;  
Best Local Similarity 28.0%; Pred. No. 2.9e-14;  
Matches 116; Conservative 46; Mismatches 130; Indels 123; Gaps 23;

Qy 23 TYRIPALIVPPAHTFLAFAEKRS--SSKDEDALHLVLRRLTQSGVOWEPLKSLMK-AT 80  
Db 86 TFRIP-LITATPRGTLTLLAFAEARKMSSDEGAKFTALRRSMDQGST--WSPATFIVNDGD 142  
Qy 81 LPHGHTMPCPVWERKSGVYVLF--ICVQGHVTERQQIMSGRNPARLCFICSDAGYSWS 139  
Db 143 VPDGLNLG-AVVSDETGVVFLFYSLCAH-----KAGCOVASTMLVWSKDDGVWS 192  
Qy 140 DVRLTEVIGPEVTHWATFVAGVGHGIGLO-----SGLIIPAYAYIIPFFWFCFLPYR 195  
Db 193 TPRNLSLD-IGTEV-----FAPGPGSGIQKQREPRKGLIIVCGHGTLEDRDGVFC----- 240  
Qy 196 ARPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIKAGHPVLYCSARTPNRRAEAL 255  
Db 241 -----LLSDDHGASWRYGS-----GVSGIP----- 260  
Qy 256 SIDHGEQFKVLSHQLECPHPGCGSVV-----SFCPLEIPGGCQDLAGEDAPAIQ 307  
Db 261 ---YGQPKQENDFPDECPYELPDGVSVINARNQNNYHCHCRIV-----LRSYDA----- 308  
Qy 308 QSPILCSSVRP-----BPE-----AGTL--SESWLLYSHTNKKRRVDLGIYLNOSPL 353  
Db 309 -----CDTLRPRDVTDFPELDPVVAAGAVVTSSGIVFFSNPAHPFVRNLI--TLRWSFS 361  
Qy 354 EAACWSRPWI-LHCGPCGYSDLAALENEGLFG-----CLFECGCKQCEQIA 399  
Db 362 NGTSWRKETVOLWPGSPGSSYSLATLEG-SMDGEGEQAPOLYVLYEKGRRNHHTESIS 415

RESULT 9  
US-09-791-537-1830  
; Sequence 1830, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1830  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-1830

Query Match 10.5%; Score 245.5; DB 5; Length 409;  
Best Local Similarity 27.1%; Pred. No. 2.2e-13;  
Matches 112; Conservative 46; Mismatches 136; Indels 119; Gaps 22;



Qy 23 TYRIPALIVPPAHTFLAFAE-KRSSKDEDAHLVLRLGLRTGOSVQWEPLKSLMKATL 81  
Db 70 TFRIP-LITATPRGTLFAFAEARKKSADGAKFTAMRS--TDGSSWSTAFIVDDGE 126  
Qy 82 PGRMTNCPVWERKSGVYLFF-TCVOGHVTERQIMSGRNPALCFCTCSODAGYSWD 140  
Db 127 ASDGLNLGAVNDVTGIVFLYITC--AHKVNCQ-----VASTMLVWSKDDGISNP 177  
Qy 141 VRLDTEEVIGPVTWATFVCGHGILQ-----SGRLIIPAYAYIIPWFECFLPYRA 196  
Db 178 PRLNSVD-IGTEM-----FAPGSGGIQKOREPGKRLVCGHGLTERDGVFC----- 224  
Qy 197 RPHSLMIYSDDLGATWHHGRKLPMTVECEVAEIVGAGHPVLYCSARTPNHRAEALS 256  
Db 225 -----LLSDDHGASWHVG-----TGVSIP----- 244  
Qy 257 IDHGEFQKPVLSHOLCEPHPGCGSVVFCPLFPGGCQDLAGED-----APAIQOS 309  
Db 245 -----FGQPKHDHF--NPDEQ-----PYELPDGSVIINARNQNNYHCRIVLRS 289  
Qy 310 PLLCSSVRP-----EPE-----ACTL--SESWLLYSHPTNKKRRVDLGIYLNOSPLEA 355  
Db 290 YDADCTLRPRDVTFOPELVDPVVAAGALATSGGIFFSNPAHPEFRVNL--TLRWSFSNG 347  
Qy 356 ACWSRPWI-LHCGPCGYSDLAALENE-----GLGCLFECGTKQCEQIA 399  
Db 348 TSWQRQVQVWPGPSYSSLTALENSTGKKQPPQLF-VLYEKGLNRYTESIS 399  
RESULT 10  
US-09-791-537-73212  
; Sequence 73212, Application us/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blonomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791.537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 73212  
; LENGTH: 605  
; TYPE: PRT  
; ORGANISM: pdb LEUT  
US-09-791-537-73212  
Query Match 8.6%; Score 201.5; DB 5; Length 605;  
Best Local Similarity 24.0%; Pred. No. 3.1e-09;  
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;  
Qy 12 PLFQOED-----KRGV-TYRIPALIVPPAHTFLAFAEKRSSKDEDAHLVL-RLGLRT 64  
Db 7 PLYTQDLAVNCREGFPNRYIPALT-VTPDGLLASYGDRPTGIDAPGPNSTLQRRSTDG 65  
Qy 65 GOSVQWEPLKSLMKATLPGRHTMPCPVWERKSGVYVLFICVQGHVTERQOIMSGRNP 124  
Db 66 GRTWGEQVVSAGQTTAPIKGFSDPYLVDRGTGTIFNF-----HVYSQRQGFAGSRPG 119  
Qy 125 -----RLCFICSDAGYSWSDVRLTTEVIGPEVTHWATFVAGPGHGILQ-----S 171  
Db 120 TDPADPNVLHANVATSTOGGLTWSH-RTITAD-ITPD-PGWRSRFAASGEGIQLRYGPHA 176  
Qy 172 GRL-----IIPAYAYIIPWFECFLPYRAPHSLMIYSDDLGATWHHGRLI-----KPMV 222  
Db 177 GRLLIOQYTIINAGAF-----QAVSVYSDDHGRWTRAGEAVGVGMNDENK 220  
Qy 223 TVECEVAEIVGKAGHPVLYCSARTPNHRAEALSIDHGECEFKPKPVLSHQLCEPPHGCCGS 282

Db 221 TVELSDGRVLLNSRD-----SAR--SGYRKVAVSTDGGHSYGPVTTIDRLDPDPTN--NAS 271  
Qy 283 VVSFPCLEIPGCGQDLAGEDAPAIQOSPLLCSSVRPEPEAGTLSWSWLLYSHTPTNKKRRV 342  
Db 272 IIRAPP-----DAP-----AGSARAKVLLFNSAASQTSR- 300  
Qy 343 DLGIYLNOSPLEAAC-----WSRPWILHCGPCGYSDLAALENEGLFGLFCGCTKQCEQ 397  
Db 301 -----SOGTIRMSCDDGQTPVSKVPFGQMSYSTLTALP-DGYTGLLYEPGT-----G 348  
Qy 398 IAFRLFTDREILSHVQDGC-----STPGMNSPESKK 428  
Db 349 IRYANFN-----LAWLGGICAPETIPDVALEPGQQ 378  
RESULT 11  
US-09-791-537-81425  
; Sequence 81425, Application us/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blonomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791.537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 81425  
; LENGTH: 647  
; TYPE: PRT  
; ORGANISM: Micromonospora viridifaciens  
US-09-791-537-81425  
Query Match 8.6%; Score 201.5; DB 5; Length 647;  
Best Local Similarity 24.0%; Pred. No. 3.4e-09;  
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;  
Qy 12 PLFQOED-----KRGV-TYRIPALIVPPAHTFLAFAEKRSSKDEDAHLVL-RLGLRT 64  
Db 49 PLYTQDLAVNCREGFPNRYIPALT-VTPDGLLASYGDRPTGIDAPGPNSTLQRRSTDG 107  
Qy 65 GOSVQWEPLKSLMKATLPGRHTMPCPVWERKSGVYVLFICVQGHVTERQOIMSGRNP 124  
Db 108 GRTWGEQVVSAGQTTAPIKGFSDPYLVDRGTGTIFNF-----HVYSQRQGFAGSRPG 161  
Qy 125 -----RLCFICSDAGYSWSDVRLTTEVIGPEVTHWATFVAGPGHGILQ-----S 171  
Db 162 TDPADPNVLHANVATSTOGGLTWSH-RTITAD-ITPD-PGWRSRFAASGEGIQLRYGPHA 218  
Qy 172 GRL-----IIPAYAYIIPWFECFLPYRAPHSLMIYSDDLGATWHHGRLI-----KPMV 222  
Db 219 GRLLIOQYTIINAGAF-----QAVSVYSDDHGRWTRAGEAVGVGMNDENK 262  
Qy 223 TVECEVAEIVGKAGHPVLYCSARTPNHRAEALSIDHGECEFKPKPVLSHQLCEPPHGCCGS 282  
Db 263 TVELSDGRVLLNSRD-----SAR--SGYRKVAVSTDGGHSYGPVTTIDRLDPDPTN--NAS 313  
Qy 283 VVSFPCLEIPGCGQDLAGEDAPAIQOSPLLCSSVRPEPEAGTLSWSWLLYSHTPTNKKRRV 342  
Db 314 IIRAPP-----DAP-----AGSARAKVLLFNSAASQTSR- 342  
Qy 343 DLGIYLNOSPLEAAC-----WSRPWILHCGPCGYSDLAALENEGLFGLFCGCTKQCEQ 397  
Db 343 -----SOGTIRMSCDDGQTPVSKVPFGQMSYSTLTALP-DGYTGLLYEPGT-----G 390  
Qy 398 IAFRLFTDREILSHVQDGC-----STPGMNSPESKK 428  
Db 391 IRYANFN-----LAWLGGICAPETIPDVALEPGQQ 420



Db 723 Y-COLSVINY-----SQRDKGDA-----IIFAIPDAN-----YNNRV 754  
Qy 337 NKRRVDL-----GYLNOSPLEACWSPWTLHCGPCGYSDLAALENEGLCFLFE-CGT 391  
Db 755 NGTVRVGLITENGSENGEPRYDIEMRYNKVAPGTGYSCLSMPN-GEIGLFYEGRGS 813  
Qy 392 KQ 393  
Db 814 RQ 815

RESULT 15  
US-09-791-537-140247  
; Sequence 140247, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 140247  
; LENGTH: 694  
; TYPE: PRT  
; ORGANISM: Clostridium perfringens  
US-09-791-537-140247

Query Match 7.3%; Score 172; DB 5; Length 694;  
Best Local Similarity 20.8%; Pred. No. 1.6e-06;  
Matches 102; Conservative 65; Mismatches 143; Indels 180; Gaps 25;  
Qy 24 YRIPALIVVPPAHTFLAFAEKRSSKD--EDALHLVLR--GLRT---GOSVQWEPLK- 74  
Db 265 YRIPALFKTEKGTLIASIDARRHGADAPNNDIDTAVRRSEDGGKTWDEGIIINDYDPS 324  
Qy 75 SLKATLPGHRTMPCPWKRSQVYLF-----FICV- 107  
Db 325 SVIDTTL-----IQDDETGRIFLLVTHPEPSKYGFNAGLGSFGKNDGKEYLCLY 374  
Qy 108 -----QGHVTER-----QQIMSGRNPAP-----LCFI 129  
Db 375 DSSGKEFTVRENVVYDKGNKTEYTNALGDLFKNGTKIDINSTAPLKAKGTSYNLV 434  
Qy 130 CSODAGYSWSDVRDLTERVIGPEVTHWATF-AVGPGHGIQLQS---CRLIIPAYAYIP 184  
Db 435 YSDDGKTWSEPNINFQV-----KKDMWKFGLGIAPGRGIQKNGEHKGRIVVPVY----- 485  
Qy 185 FWFFCFRLPY---RAPHSLMIYSDDLGATW-----HHGRLIK-----PM 221  
Db 486 -----YTNEKGKSSAVIYSDDSKKNWTIGESPNDNRKLENGKIINSKTLSDDAPO 536  
Qy 222 VTVECEVAETGKAGHPVLYCSARTPNHRAEALSIDHGECFQKPVLSHQICEPPHQCQG 281  
Db 537 LT-ECQVEM--PNGQLKLF--MRNLGYLNIATSFDDGATWDETVEKDTNVLEPY-CQL 590  
Qy 282 SVWSFCPLIEPGGCDLAGEDAPAIQOSPILCSSVRPEPEAGTLESWLLYSHPTNKKR- 340  
Db 591 SVINY-----SOKVDGKA-----VIFSNPNARSRS 616  
Qy 341 ----RVDL-----GYLNOSPLEACWSPWTLHCGPCGYSDLAALENEGLCFLFECGTK 392  
Db 617 NGTVRIGLINQVGYENGEPKPYEDWKYNKLVKPGYAYSCLTSLN-GNIGLLYEGTPS 675  
Qy 393 QCEQIAFRL 402  
Db 676 EEMSYIEMNL 685

Search completed: October 7, 2002, 15:05:57  
Job time: 439 sec

